

Plasmodium falciparum; circumsporozoite; anti-malarial; CTP; cholesterol ester transport protein; anti-arteriosclerotic. Chimeric - Measles virus. Chimeric - Rattus sp. MO9966957-A2. 29-DEC-1999. 21-JUN-1999; 99WO-0513975. 20-JUN-1998; 98US-0100412. (UNB1-) UNITED BIOMEDICAL INC. Wang CY; WPI; 2000-160564/14. The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes of the invention along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response, specifically against Plasmodium falciparum, cholesterol ester transport protein (CTP) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of interleukin hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration); for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AA91121 represents a promiscuous T helper epitope from the measles virus F (MVF) protein and sequences AA91122-Y91142, AA91126 and AA91245-Y91246 represent synthetic Th epitopes based on the MVF Th epitope. Sequence AA91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences AA91144-Y91115 are synthetic epitopes derived from this HBV epitope. AA91156-Y91156, AA91227 and AA91247-Y91247 are antigenic peptides comprising an LHRH sequence, joined to a promiscuous Th epitope. AA91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AA91200 is somatostatin, and AA91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AA91208 is a human CD4 CDR2-like domain antigenic site, and AA91209-Y90211 are MWN Th epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. AA990212 is a modified version of a human IGE (Immunoglobulin E) CH3 domain, and AA990213-Y90219 are Th epitopes. CC antigenic peptides which may be used in the treatment of allergies. AA91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AA91221-Y91222 comprise this peptide and a Th epitope. AA91223 is a Plasmodium falciparum circumsporozoite (CS) target epitope, and AA91224-Y91225 comprise the CS antigen and an MVF Th antigen, and AA91224-Y91225 comprise the CS antigen and an MVF Th epitope, and may be used in a malaria vaccine. AA91228-Y91231 represent CTP-derived peptides and AA91232-Y91241 are immunogens comprising a CTP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AA91247 and AA91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AA91248-Y91251 and AA91256-Y91273 are antigenic peptides comprising MWN Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AA91198 and AA91199 are respectively an immunostimulatory Invasin

CC	protein/peptide from <i>Yersinia</i> species, and hinge spacer peptide, both of
CC	which may optionally be used in the antigenic peptides of the
CC	invention.
CC	
XX	
XX	
Sequence	28 AA:
Query Match	42.7%; Score 106; DB 21; Length 28;
Best Local Similarity	64.3%; Pred. NO.4.2e-07;
Matches 18; Conservative	6; Mismatches 2; Indels 2; Gaps 1;
Oy	19 LSEIKGVIVHRLGVEGSPSLHMSYGLRP 46
Db	2 ISEIRGIITHRIEGIGGE--HMSYGLRP 27

```

RESULT 15
AAW67581
ID AAW67581 standard; peptide: 40 AA.
XX
AC AAW67581;
XX
DT 02-MAR-1999 (first entry)
XX
DE Synthetic chimera fimbriin/T-cell epitope peptide LBL.
XX
KM Chimeric; non-tyrable Haemophilus influenzae; fimbriin: T-cell epitope;
KW immunogenic composition; immune response.
XX
OS Synthetic.
PN US5843464-A.
XX
PD 01-DEC-1998.
XX
PF 02-JUN-1995; 95US-0460502.
XX
PR 02-JUN-1995; 95US-0460502.
PA (OHIS ) UNIV OHIO STATE.
PI Bakaletz IO, Kaumaya PNP;
XX
DR WPI: 1999-044514/04.
XX
PT Synthetic chimeric fimbriin peptide - useful for vaccination against
PT non-tyrable Haemophilus influenzae
PS
PS Claim 4; Column 4; 16pp; English.
XX
CC The invention relates to the manufacture of a synthetic chimeric peptide
CC comprising a non-tyrable Haemophilus influenzae fimbriin peptide fused via
CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is
CC used in immunogenic compositions which induce an immune response against
CC non-tyrable Haemophilus influenzae. This sequence represents an example
CC of a chimeric fimbriin/T-cell epitope peptide and is designated LBL.
CC The peptide comprises a 19 amino acid sequence corresponding to amino
CC acids 117-135 of the fimbriin protein, the linker sequence and amino acid
CC 288-302 of the measles virus fusion protein (a T-cell epitope).
CC
SQ Sequence 40 AA:
QY Query Match 42.7%; Score 106; DB 20; Length 40;
Best Local Similarity 79.3%; Pred. No. 6, 4e-07;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0.
QY 6 GLRPGSSGPSKLKLSIKGYIYHRLGVE 34
| | ||||| |||||||||
12 GTRDHRKGPSLKLKLSIKGYIYHRLGVE 40

```

CC invasin domain immunostimulatory peptide of *Yersinia sp.*, a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The synthetic Th epitope is derived
 CC from a structured synthetic library (SSAL) designated SSAL Th1.
 CC SSAL Th1 is modeled after a promiscuous epitope taken from the F protein
 CC of the Measles virus. The peptide immunogens cause induction of a
 CC specific immune response to LHRH which is involved in regulation of
 CC spermatogenesis, ovulation, oestrus, sexual development and secretion
 CC of sex hormones. Provision of a promiscuous T helper epitope (which is
 CC functional in genetically diverse subjects) provides optimum
 CC immunogenicity to the B cell epitopes of the target antigen and thus
 CC high antibody titres against the target antigen. The peptide immunogens
 CC of the invention are used to vaccinate against mammalian LHRH, for use
 CC as (reversible) contraceptive; control of hormone-dependent tumours
 CC (cancer of prostate or breast, also endometriosis); to prevent boar
 CC taint (and improve meat quality) and for immunocastration.

SQ Sequence 45 AA;

Query Match 43.5%; Score 108; DB 21; Length 45;
 Best Local Similarity 71.4%; Pred. No. 4e-07; Mismatches 2; Gaps 1;
 Matches 20; Conservative 4; Indels 2;

OY 19 LSEIKGIVHRLEGEVGPSTLHMSYGLRP 46
 :|||||||:|:| |
 Db 19 ISEIKGVIVHKIEGIGGE--HMSYGLRP 44

RESULT 13
 ID AAY91179 standard; peptide: 31 AA.
 XX AAY91179;
 XX 22-MAY-2000 (first entry)
 DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:59.
 XX
 XX Promiscuous T-cell epitope: measles virus F protein; MVF.
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
 KW Plasmodium falciparum; circumporozoite; antimalarial; CERP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.
 XX
 OS Chimeric - Measles virus.
 OS Chimeric - Rattus sp.
 OS
 PN MO9966957-A2.
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99MO-US13975.
 PR 20-JUN-1998; 98US-0100412.
 XX
 PA (UNBT-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 DR WPI; 2000-160564/14.
 XX
 PT New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus
 XX
 PS Example 1; Page 86; 12pp; English.
 XX
 CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,

CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91126 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVF Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to protect HIV
 CC infection of T cells. AAY90212 is a modified version of a human IGE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVF Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from *Yersinia species*, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 CC
 XX
 XX Sequence 31 AA;
 SQ

Query Match 43.1%; Score 107; DB 21; Length 31;
 Best Local Similarity 75.0%; Pred. No. 3.5e-07;
 Matches 21; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLEGEVGPSTLHMSYGLRP 46
 :|||||||:|:| |
 Db 3 LSEIKGVIVHKIEGMLFGGEHMSYGLRP 30

RESULT 14
 ID AAY91158 standard; peptide: 28 AA.
 XX AAY91158;
 XX 22-MAY-2000 (first entry)
 DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:38.
 XX
 XX Promiscuous T-cell epitope: measles virus F protein; MVF.
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
 KW

KM luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KM foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KM Plasmodium falciparum; circumsporozoite; antimalarial; CEMP;
 KM cholesterol ester transport protein; anti-arteriosclerotic.
 XX Chimeric - Measles virus.
 OS Chimeric - Rattus sp.
 XX MO9966957-A2.
 PN 29-DEC-1999.
 PD 29-DEC-1999.
 XX 21-JUN-1999; 99MO-US13975.
 XX 20-JUN-1998; 98US-0100412.
 PR (UNBI-) UNITED BIOMEDICAL INC.
 PA Wang CY;
 PI WPI; 2000-160564/14.
 DR New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus
 XX Example 1; Page 80; 129pp; English.
 PS
 XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CEMP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IGE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CEMP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CEMP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and

CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 XX
 XX
 XX Sequence 45 AA;
 SQ
 Query Match 43.5%; Score 108; DB 21; Length 45;
 Best local Similarity 71.4%; Pred. No. 4e-07;
 Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
 OY 19 LSEIKGYVHRLGEGSPSLHWSYGLRP 46
 Db 19 ISEIKGYVHKIEIGEG--HWSYGLRP 44
 :|||||:|||||
 RESULT 12
 AAY68573
 ID AAY68573 standard; peptide; 45 AA.
 XX
 XX AAY68573;
 AC
 XX 05-MAY-2000 (first entry)
 DT
 XX Peptide immunogen comprising a Th epitope and LHRH target antigen.
 DE
 XX
 XX Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
 KM luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KM oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KM vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KM breast cancer; endometriosis; boar taint; meat quality;
 KM invasiv domain; immunocastration.
 XX
 XX Synthetic.
 OS Yersinia sp.
 OS Measles virus.
 OS Unidentified.
 OS
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..16
 FT /note= "Invasiv domain AAY68565"
 FT Peptide 17..18
 FT /note= "spacer"
 FT Peptide 19..33
 FT /note= "helper Th epitope AAY68544"
 FT Peptide 34..35
 FT /note= "spacer"
 FT Peptide 36..45
 FT /note= "LHRH antigenic epitope AAY68566"
 XX
 PN MO9966952-A1.
 XX 29-DEC-1999.
 PD 29-DEC-1999.
 XX 21-JUN-1999; 99MO-US13960.
 XX 20-JUN-1998; 98US-0100414.
 PR (UNBI-) UNITED BIOMEDICAL INC.
 PA Wang CY;
 PI WPI; 2000-160564/14.
 DR New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer
 XX
 PS Claim 9; Page 71; 102pp; English.
 XX The present sequence represents a peptide immunogen comprising an

```
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IGE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CERP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MWF Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.
XX
SO Sequence 27 AA:
Query Match 43.5%; Score 108; DB 21; Length 27;
Best Local Similarity 71.4%; Pred. No. 2.2e-07;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
OY 19 LSEIKGVIVHRLGVESPSLHMSYGLRP 46
:|||||:|||||:|||||:|||||:|||||:
DB 1 MSEIKGVIVHRLKGIGE--HMSYGLRP 26
RESULT 10
AAY91167
ID AAY91167 standard; peptide: 27 AA.
XX
AC AAY91167;
DT 22-MAY-2000 (first entry)
XX
DE Modified MWF Th epitope/LHRH antigenic peptide, SEQ ID NO:47.
XX
XX Promiscuous T-cell epitope; measles virus F protein; MWF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW interleukin hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
KW cholesterol ester transport protein; anti-arteriosclerotic.
XX
XX Chimeric - Measles virus.
OS Chimeric - Rattus sp.
OS
XX
PN WO966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99MO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
XX New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus
PS Example 1; Page 81; 129pp; English.
XX
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
```

```
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of interleukin hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration; for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MWF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MWF Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IGE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MWF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CERP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MWF Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.
XX
SO Sequence 27 AA:
Query Match 43.5%; Score 108; DB 21; Length 27;
Best Local Similarity 75.0%; Pred. No. 2.2e-07;
Matches 21; Conservative 3; Mismatches 2; Indels 2; Gaps 1;
OY 19 LSEIKGVIVHRLGVESPSLHMSYGLRP 46
:|||||:|||||:|||||:|||||:|||||:
DB 1 MSEIKGVIVHRLKGIGE--HMSYGLRP 26
RESULT 11
AAY91165
ID AAY91165 standard; peptide: 45 AA.
XX
AC AAY91165;
DT 22-MAY-2000 (first entry)
XX
DE Modified MWF Th epitope/LHRH antigenic peptide, SEQ ID NO:45.
XX
XX Promiscuous T-cell epitope; measles virus F protein; MWF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
```


CC		recurrent functional ovarian cysts, (severe) premenstrual syndrome or oestrogen-dependent breast cancer, or for induction of infertility.
XX		
SO	Sequence	45 AA:
	Query Match	46.8%; Score 116; DB 15; Length 45;
	Best Local Similarity	85.7%; Pred. No.3.5e-08;
	Matches 24:	Conservative 0; Mismatches 2; Gaps 1;
OY	19 LSEIKGVIVHRLEGVEGPSLHMSYGLRP 46 19 LSEIKGVIVHRLEGVGE--HMSYGLRP 44	
DB		
RESULT 7		
ID	AAAY91163	
AC	AAAY91163 standard; peptide; 27 AA.	
XX	AAAY91163;	
DT	22-MAY-2000 (first entry)	
DE	Modified MYF Th epitope/LHRH antigenic peptide, SEQ ID NO:43.	
XX		
KM	Promiscuous T-cell epitope; measles virus F protein; MYF.	
KM	hepatitis B virus surface antigen; HBV; Immunogenic; B-cell epitope;	
KM	luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;	
KW	somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;	
KW	foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;	
KM	Plasmodium falciparum; circumsporozoite; antimalarial; CEMP;	
KM	cholesterol ester transport protein; anti-arteriosclerotic.	
XX		
OS	Chimeric - Measles virus.	
OS	Chimeric - Rattus sp.	
PM	MO9966957-A2.	
PD	29-DEC-1999.	
XX		
PF	21-JUN-1999; 99WO-USJ3975.	
XX		
PR	20-JUN-1998; 98US-0100412.	
PA	(UNBI-) UNITED BIOMEDICAL INC.	
XX		
PI	Wang CY;	
DR	WPI: 2000-160564/14.	
XX		
PT	New artificial T helper cell epitope and derived immunogens with target antigenic site, for immunization against e.g. malaria, arteriosclerosis or human immune deficiency virus	
PT	-	
PS	Example 1; Page 80; 129pp; English.	
XX		
CC	The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes of the invention along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response.	
CC	specifically against Plasmodium falciparum, cholesterol ester transport protein (CEP) or HIV epitopes, but more generally against any pathogen,	
CC	Immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone-	
CC	dependent cancer, prevention of boar taint in meat, and	
CC	immunocastration); for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper	
CC		

Sequence	27 AA:
Query Match	45.6%; Score 113; DB 21; Length 27;
Best Local Similarity	82.1%; Pred. NO. 4.7e-08;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1.	
Oy	19 LSEIKGVIVHRLEGEGPSLHMSYGLRP 46 IIIIIIIIII I IIIIII
Db	1 LSEIKGVIVHRLEGEGGE--HMSYGLRP 26
RESULT 8	
AAAY91175	
XX	AAAY91175 standard; peptide: 31 AA.
XX	
DT	22-MAY-2000 (first entry)
XX	
DE	Modified MVE Th epitope/LHRH antigenic peptide, SEQ ID NO:55.
XX	
KW	Promiscuous T-cell epitope: measles virus F protein; MVE.
KW	hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW	luteinizing hormone releasing hormone; LHRH; contraceptive; anticancer;
KW	somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMVY;
KW	foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW	Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
KW	cholesterol ester transport protein; anti-arteriosclerotic.
XX	
OS	Chimeric - Measles virus.
OS	Chimeric - Rattus sp.
FN	
XX	WO9966957-A2.
XX	
PD	29-DEC-1999.
XX	
FE	21-JUN-1999; 99WO-US13975.
XX	
XX	20-JUN-1998; 98US-0100412.
XX	

Best Local Similarity	85.7%	Pred. No. 1.9e+08	
Matches	24	Conservative	0
		Mismatches	2
		Indels	2
		Gaps	1

XX	RESULT 5
XX	AAV68567
ID	AAV68567 standard; peptide; 27 AA.
XX	
AC	AAV68567;
DT	05-MAY-2000 (first entry)
XX	
DE	Peptide immunogen comprising a Th epitope and LHRH target antigen.
XX	
KW	Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH
KW	luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW	oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW	vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW	breast cancer; endometriosis; boar taint; meat quality; chimera;
XX	immunocastration.
OS	Chimeric - Measles virus.
OS	Chimeric - Unidentified.
XX	
FH	Key
FT	Location/Qualifiers
FT	Peptide
FT	1..15
FT	/note-"helper Th epitope AAV68540"
FT	Peptide
FT	16..17
FT	/note-"spacer"
FT	Peptide
FT	18..27
FT	/note-"LHRH antigenic epitope AAV68566"
XX	
PN	MO9966952-A1.
XX	
PD	29-DEC-1999.
XX	
PP	21-JUN-1999; 99MO-US13960.
PR	20-JUN-1998; 98US-0100414.
XX	
PA	(UNBI-) UNITED BIOMEDICAL INC.
P1	Wang CY;
PI	
PT	New peptide immunogen containing luteinising hormone-releasing hormone
PT	antigen site and helper T cell epitope, for e.g. contraception and
PT	treatment of cancer -
PS	
PS	Example 1; Page 63; 102pp; English.
XX	
CC	The present sequence represents a peptide immunogen comprising a
CC	helper T cell (Th) epitope of the F protein of the Measles virus and
CC	a target antigen, luteinising hormone-releasing hormone (LHRH).
CC	The peptide immunogens cause induction of a specific immune response
CC	to LHRH which is involved in regulation of spermatogenesis, ovulation,
CC	oestrus, sexual development and secretion of sex hormones. Provision of
CC	a promiscuous T helper epitope (which is functional in genetically
CC	diverse subjects) provides optimum immunogenicity to the B cell
CC	epitopes of the target antigen and thus high antibody titres against
CC	the target antigen. The peptide immunogens of the invention are used
CC	to vaccinate against mammalian LHRH, for use as (reversible)
CC	contraceptive; control of hormone-dependent tumours (cancer of prostate
CC	or breast, also endometriosis); to prevent boar taint (and improve meat
CC	quality) and for immunocastration.
XX	
SQ	Sequence 27 AA;

Query Match	46.8%	Score 116:	DB 21	Length 27:
Best Local Similarity	85.7%	Pred. No. 1.9e+08:		
Matches 24:	Conservative	0:	Mismatches 2:	Indels 2:
Gaps	1:			
QY	19	LSEIKGVIVHRLGEGPSLHMSYGLRP	46	
Db	1	LSEIKGVIVHRLGEGVGE--HMSYGLRP	26	

RESULT 6
AAR62721

ID AAR62721 standard; peptide: 45 AA.
AC AAR62721;
XX
DT 10-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
KW Helper T cell epitope; universal immune stimulator; invasive; hapten;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
RW androgen-dependent carcinoma; antitumour; infertility;
KW measles virus F protein.
XX
OS Synthetic.

	Key	Location/Qualifiers
FH	Domain	1..16
FT	/note= "Invasin domain"	
FT	Domain	19..33
FT	/note= "measles virus F protein helper T cell epitope"	
FT	Domain	36..45
FT	/note= "LHRH hapten"	
PN	WO9425060-A.	
PD	10-NOV-1994.	
PF	28-APR-1994;	94WO-US04832.
PR	27-APR-1993;	93US-0057166.
PR	14-APR-1994;	94US-0229275.
PA	(IADD/) LADD A E.	
PA	(WANG/) WANG C Y.	
PA	(ZAMB/) ZAMB T.	
PI	Ladd AE, Wang CY, Zamb T;	
DR	WPI: 1994-357910/44.	
PT	Immunogenic luteinising hormone releasing hormone peptide(s) - that suppress LHRH activity in males and females	

Claim 8: Page 88; 213pp: English.

Synthetic immunogenic peptides are provided in which a universal immune stimulator is linked to a peptide or protein hapten containing B cell and/or cytotoxic T lymphocyte epitopes, giving a product which causes potent immune responses to the coupled peptide or protein. The stimulator consists of (A) a promiscuous helper T cell epitope (Th) which elicits an immune response to the coupled peptide in members of a heterogeneous population expressing diverse HLA phenotypes, and (B) an adjuvant amino acid sequences (e.g. Gly-Gly) can be provided between the invasin and Th domains and between the immune stimulator and hapten components. When the hapten is LHRH, then optionally the invasin domain can be omitted from the immune stimulator component.

The present sequence represents an LHRH-containing immunogenic peptide as above which can be used as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-dependent carcinoma, prostatic carcinoma, testicular carcinoma, endometriosis, benign uterine tumours,

XX 28-APR-1994; 94WO-US04832.
XX
XX 27-APR-1993; 93US-0057166.
XX
XX 14-APR-1994; 94US-0229275.
XX
XX (LADD/) LADD A E.
XX (WANG/) WANG C Y.
XX (ZAMB/) ZAMB T.
XX
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
XX that suppress LHRH activity in males and females
XX
XX Claims 8, 12; Page 86; 213pp; English.
XX
XX Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein hapten containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The
XX stimulator consists of (A) a promiscuous helper T cell epitope (Th)
XX which elicits an immune response to the coupled peptide in members of
XX a heterogeneous population expressing diverse HLA phenotypes, and (B)
XX an adjuvant peptide sequence from the invasins protein of *Yersinia*.
XX Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
XX invasins and Th domains and between the immune stimulator and hapten
XX components. When the hapten is LHRH, then optionally the invasins domain
XX can be omitted from the immune stimulator component.
XX The present sequence represents an LHRH-containing, invasins-free
XX immunogenic peptide as above which can be used as a potent vaccine for
XX treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
XX prostatic carcinoma, testicular carcinoma, endometriosis, benign
XX uterine tumours, recurrent functional ovarian cysts, (severe)
XX premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX induction of infertility.
XX This sequence is particularly preferred.
XX
XX Sequence 27 AA:
XX
XX Query Match 46.8%; Score 116; DB 15; Length 27;
XX Best Local Similarity 85.7%; Pred. No. 1.9e-08;
XX Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
XX
XX 19 LSEIKGVYHRLGEGSPSLHWSYGLRP 46
XX ||||||||||||||| |||||||
XX Db 1 LSEIKGVYHRLGEGGGE--HWSYGLRP 26
XX
XX RESULT 4
XX AAY91156
XX ID AAY91156 standard; peptide; 27 AA.
XX
XX AAY91156:
XX
XX 22-MAY-2000 (first entry)
XX
XX MWF Th epitope/LHRH antigenic peptide, SEQ ID NO:36.
XX
XX Promiscuous T-cell epitope; measles virus F protein; MVR;
XX hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
XX luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
XX somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
XX foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
XX Plasmodium falciparum; circumsporozoite; antimalarial; CEMP;
XX cholesterol ester transport protein; anti-arteriosclerotic.
XX
XX Chimeric - Measles virus.
XX Chimeric - Rattus sp.
XX
XX MO9966957-A2.
XX
XX

XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13975.
XX
XX 20-JUN-1998; 98US-0100412.
XX
XX (UNBT-) UNITED BIOMEDICAL INC.
XX
XX Wang CY;
XX
XX WPI; 2000-160564/14.
XX
XX New artificial T helper cell epitope and derived immunogens with target
XX antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX or human immune deficiency virus -
XX
XX Example 1; Page 77; 129pp; English.
XX
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
XX and immunogenic peptides comprising the Th epitopes of the invention
XX along with B cell epitopes. The Th epitopes and peptide immunogens
XX containing them, are used to induce a T helper cell response,
XX specifically against Plasmodium falciparum, cholesterol ester transport
XX protein (CEMP) or HIV epitopes, but more generally against any pathogen,
XX immunoreactive self-antigen or tumour antigen. The Th epitopes and
XX peptide immunogens may be used for prevention and/or treatment of
XX infections (HIV, foot-and-mouth disease or malaria); for cancer
XX immunotherapy; for inhibition of the action of luteinising hormone
XX releasing hormone (LHRH) for contraception, treatment of hormone-
XX dependent cancer, prevention of boar taint in meat, and
XX immunocastration; for promoting the growth of animals; or for
XX treating allergies or arteriosclerosis. Incorporation of a promiscuous
XX Th (functional in genetically diverse subjects) into an immunogen
XX improves capacity to induce a strong T helper cell-mediated immune
XX response, resulting in production of antibodies against a target
XX antigen. Th can replace carrier proteins and pathogen-derived T helper
XX epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
XX from the measles virus F (MVR) protein and sequences AAY91122-Y91142,
XX AAY9126 and AAY91245-Y91246 represent synthetic Th epitopes based on the
XX MVR Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
XX from hepatitis B virus (HBV) surface antigen, and sequences
XX AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
XX AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides.
XX comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
XX is the LHRH target antigenic peptide used in these LHRH antigenic
XX peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
XX peptides comprising somatostatin and a Th epitope. Somatostatin
XX immunogens may be used to promote growth in livestock. AAY91208 is a
XX human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVR Th
XX epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
XX infection of T cells. AAY90212 is a modified version of a human IGE
XX (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3
XX antigenic peptides which may be used in the treatment of allergies.
XX AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
XX VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
XX epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
XX antigen, and AAY91224-Y91225 comprise the CS antigen and an MVR Th
XX epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
XX CEMP-derived peptides and AAY91232-Y91241 are immunogens comprising a
XX CEMP peptide and a Th epitope which may be used to prevent or treat
XX arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
XX are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
XX AAY91258-Y91273 are antigenic peptides comprising MVR Th and HIV-1 B-cell
XX epitope which may be used as a component in an anti-HIV-1 vaccine.
XX AAY91198 and AAY91199 are respectively an immunostimulatory invasins
XX protein epitope from *Yersinia* species, and hinge spacer peptide, both of
XX which may optionally be used in the antigenic peptides of the
XX invention.
XX
XX Sequence 27 AA:
XX
XX Query Match 46.8%; Score 116; DB 21; Length 27;
XX
XX

PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 DR Immunogenic luteinising hormone releasing hormone peptide(s) -
 XX PT that suppress LHRH activity in males and females
 XX
 XX Claim 8; Page 84; 213pp; English.
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC Invasin and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the Invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, Invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX
 XX Sequence 25 AA:
 SQ
 Query Match 47.2%; Score 117; DB 15; Length 25;
 Best Local Similarity 85.7%; Pred. No. 1.3e-08;
 Matches 24; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 QY 19 LSEIKGYIVHRLGEGPSLHWSYGLRP 46
 DB 1 LSEIKGYIVHRLGEGV---HWSYGLRP 24
 RESULT 2
 AAR62708
 ID AAR62708 standard; peptide; 42 AA.
 AC AAR62708;
 XX
 DT 10-SEP-1995 (first entry)
 XX
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; Invasin; hapten;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 XX Synthetic.
 OS
 XX
 FT Key Location/Qualifiers
 FT Domain 1..15
 FT /note- "measles virus F protein helper T cell epitope"
 FT Domain 16..30
 FT /note- "measles virus F protein helper T cell epitope"
 FT Domain 33..42
 FT /note- "LHRH hapten"
 XX
 PN W09425060-A.
 XX
 XX 10-NOV-1994.
 PD
 XX 28-APR-1994; 94WO-US04832.

PR 27-APR-1993; 93US-0057166.
 PR 14-APR-1994; 94US-0229275.
 XX
 PA (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 DR Immunogenic luteinising hormone releasing hormone peptide(s) -
 XX PT that suppress LHRH activity in males and females
 XX
 XX Claim 8; Page 86; 213pp; English.
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC Invasin and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the Invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, Invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX
 XX Sequence 42 AA:
 SQ
 Query Match 47.2%; Score 117; DB 15; Length 42;
 Best Local Similarity 82.8%; Pred. No. 2.4e-08;
 Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
 QY 18 LSEIKGYIVHRLGEGPSLHWSYGLRP 46
 DB 15 VLSEIKGYIVHRLGEGVGE--HWSYGLRP 41
 RESULT 3
 AAR62707
 ID AAR62707 standard; peptide; 27 AA.
 AC AAR62707;
 XX
 DT 10-SEP-1995 (first entry)
 XX
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; Invasin; hapten;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 XX Synthetic.
 OS
 XX
 FT Key Location/Qualifiers
 FT Domain 1..15
 FT /note- "measles virus F protein helper T cell epitope"
 FT Domain 18..27
 FT /note- "LHRH hapten"
 XX
 PN W09425060-A.
 XX
 XX 10-NOV-1994.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 24.7051 Seconds
(without alignments)
211.311 Million cell updates/sec

Title: us-09-848-834A-17

Perfect score: 248
Sequence: 1 XHMSYGLRPSSSGPSLKLKLS.....HRLGVEGPSLHMSYGLRPX 47

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	47.2	25	15	AAK62705
2	117	47.2	42	15	AAK62708
3	116	46.8	27	15	AAK62707
4	116	46.8	27	21	AAV91156
5	116	46.8	27	21	AAV68567
6	116	46.8	45	15	AAK62721
7	113	45.6	27	21	AAV91163
8	110	44.4	31	21	AAV91175
9	108	43.5	27	21	AAV91161
10	108	43.5	27	21	AAV91167
11	108	43.5	45	21	AAV91165

12	108	43.5	45	21	AAV68573
13	107	43.1	31	21	AAV91179
14	106	42.7	28	21	AAV91158
15	106	42.7	40	20	AAW67581
16	106	42.7	40	21	AAV79986
17	105	42.3	31	21	AAV91173
18	105	42.3	31	21	AAV68582
19	105	42.3	47	21	AAV91180
20	105	42.3	47	21	AAV68583
21	105	42.3	49	21	AAV91177
22	103	41.5	27	21	AAV91170
23	103	41.5	27	21	AAV68575
24	103	41.5	35	21	AAV91242
25	101	40.7	40	20	AAW67582
26	100	40.3	28	15	AAK62726
27	100	40.3	46	15	AAK62728
28	100	40.3	47	21	AAV91183
29	100	40.3	47	21	AAV68586
30	99	39.9	28	21	AAV91159
31	96.5	38.9	42	21	AAK20865
32	96.5	38.9	49	17	AAW03944
33	96.5	38.9	49	19	AAW79567
34	96.5	38.9	49	19	AAW61542
35	96.5	38.9	49	21	AAV58363
36	96.5	38.9	49	21	AAV58135
37	96.5	38.9	544	17	AAW03943
38	96.5	38.9	544	19	AAW79570
39	96.5	38.9	695	19	AAW79573
40	96.5	38.9	695	21	AAV58361
41	96.5	38.9	695	21	AAV58133
42	96.5	38.9	977	17	AAW03942
43	96.5	38.9	977	19	AAW79569
44	94.5	38.1	30	11	AAK07323
45	94.5	38.1	40	20	AAV31183

ALIGNMENTS

RESULT 1					
ID	AAK62705	standard; peptide; 25 AA.			
AC	AAK62705;				
DT	10-SEP-1995	(first entry)			
DE	LHRH-containing immunogenic peptide.				
KW	Helper T cell epitope; universal immune stimulator; invasive; hapten; vaccine; LHRH; interleukin hormone releasing hormone; prostate; androgen-dependent carcinoma; antitumour; infertility; measles virus F protein.				
OS	Synthetic.				
FT	Key	Location/Qualifiers			
FT	Domain	1..15			
FT	Domain	/note="measles virus F protein helper T cell epitope"			
FT	Domain	16..25			
FT	Domain	/note="LHRH hapten"			
XX	W09425060-A.				
XX	10-NOV-1994.				
XX	28-APR-1994;	94WO-US04832.			
XX	27-APR-1993;	93US-0057166.			
XX	14-APR-1994;	94US-0229275.			
XX	(LADD/)	LADD A E.			
XX	(WANG/)	WANG C Y.			

Peptide immunogen
Modified MVF Th ep
Modified MVF Th ep
Synthetic chimera f
Measles virus fusi
Modified MVF Th ep
Peptide immunogen
Inv epitope/modifi
Modified immunogen
Modified MVF Th ep
Modified MVF Th ep
Peptide immunogen
Modified MVF Th ep
Synthetic chimera f
LHRH-containing im
Inv epitope/modifi
Peptide immunogen
Modified MVF Th ep
GnRH tandem dimer
GnRH 4-repeat sequ
GnRH-2. Synthetic
Peptide hormone gn
Four-copy gonadotr
GnRH analogue mult
LKT-GnRH protein f
LKT-GnRH chimeric
LKT-GnRH chimeric
Leukotoxin/gonadot
Gonadotropin relea
LKT-GnRH protein f
LKT-GnRH chimeric
Interleukin hormon
Ubiquitin fusion p

AC 030EW7
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE FUSION PROTEIN.
 OS Measles virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
 OX NCBI_TaxID=11234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OSA-2;
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
 RA Ogura H.;
 RT "Nucleotide sequences of the fusion protein gene of subacute
 sclerosing panencephalitis viruses: deduced amino acid sequences
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or
 RT predicted secondary structure changed.";
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF179438; AF02703.1; -.
 DR HSSP; P04849; ISVF.
 DR InterPro: IPR000776; Fusion_gly.
 DR Pfam: PF00523; fusion_gly; 1.
 SQ SEQUENCE 550 AA; 59333 MW; 086E51FED582BBA CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVIYHRLGV 33
 |||||
 DB 288 LSEIKGVIYHRLGV 302

RESULT 14
 O9WMK4
 ID O9WMK4 PRELIMINARY; PRT; 550 AA.
 AC O9WMK4;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE FUSION PROTEIN.
 GN F.
 OS Measles virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
 OX NCBI_TaxID=11234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WTF;
 RX MEDLINE-99329215; PubMed-10400788;
 RA Johnston I.C., Ter Meulen V., Schneider-Schaulies J.,
 RA Schneider-Schaulies S.;
 RT "A recombinant measles vaccine virus expressing wild-type
 RT glycoproteins : consequences for viral spread and cell tropism.";
 RL J. Virol. 73:6903-6915(1999).
 DR EMBL: AJ133108; CAB38075.1; -.
 DR HSSP; P04849; ISVF.
 DR InterPro: IPR000776; Fusion_gly.
 DR Pfam: PF00523; fusion_gly; 1.
 SQ SEQUENCE 550 AA; 59580 MW; 825549968B5D862 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVIYHRLGV 33
 |||||
 DB 288 LSEIKGVIYHRLGV 302

RESULT 15
 P90330

ID P90330 PRELIMINARY; PRT; 550 AA.
 AC P90330;
 DT 01-MAY-1997 (TReMBLrel. 03, Created)
 DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE FUSION PROTEIN.
 GN F.
 OS Measles virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
 OX NCBI_TaxID=11234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NAGAHATA(HB);
 RA Sheng J., Watanabe M., Ueda S.;
 RT "Selection of a neurotropic variant of measles virus.";
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: D63924; BAA09951.1; -.
 DR HSSP; P04849; ISVF.
 DR InterPro: IPR000776; Fusion_gly.
 DR Pfam: PF00523; fusion_gly; 1.
 SQ SEQUENCE 550 AA; 59589 MW; 73E7BD457ABA39B7 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVIYHRLGV 33
 |||||
 DB 288 LSEIKGVIYHRLGV 302

Search completed: October 10, 2002, 16:10:05
 Job time : 19.2778 secs

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MASUSAKO:
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF179430; AAF02695.1; -.
DR HSSP: P04849; ISVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly.1.
SO SEQUENCE 550 AA; 59559 MW; 609EE024A7E59C4 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLEGV 33
DB 288 LSEIKGIVHRLEGV 302

RESULT 10

O9QEX0 PRELIMINARY; PRT; 550 AA.
AC O9QEX0:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TOYOSHIMA:
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF179432; AAF02697.1; -.
DR HSSP: P04849; ISVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly.1.
SO SEQUENCE 550 AA; 59504 MW; 2AA969D37FA5CA17 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLEGV 33
DB 288 LSEIKGIVHRLEGV 302

RESULT 11
O9QEW9 PRELIMINARY; PRT; 550 AA.
ID O9QEW9:
AC O9QEW9;

DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OSA-2:
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF179436; AAF02701.1; -.
DR HSSP: P04849; ISVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly.1.
SO SEQUENCE 550 AA; 59405 MW; 0AE5DBFC5DD22BBA CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLEGV 33
DB 288 LSEIKGIVHRLEGV 302

RESULT 12

O9QEW8 PRELIMINARY; PRT; 550 AA.
AC O9QEW8:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OSA-2:
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF179437; AAF02702.1; -.
DR HSSP: P04849; ISVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly.1.
SO SEQUENCE 550 AA; 59315 MW; 086E51FED235EBBA CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLEGV 33
DB 288 LSEIKGIVHRLEGV 302

RESULT 13
O9QEW7 PRELIMINARY; PRT; 550 AA.
ID O9QEW7:
AC O9QEW7;

RA Ayanov R.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA Gusev A.A.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY035887; AAK63190.1; -
SQ SEQUENCE 546 AA; 58572 MW; 449B2BD7405F0B CRC64;

Query Match 29.0%; Score 72; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLGV 33
Db 284 LSEIKGIVHRLGV 298
|||||

RESULT 6
OY 089495 PRELIMINARY; PRT; 550 AA.
AC 089495;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92230209; PubMed-1566568;
RA Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;
RT "Genetic variability of the glycoprotein genes of current wild-type
RT measles isolates."
RL Virology 188:135-142(1992).
DR EMBL: M81903; AAA6422.1; -
DR EMBL: M81901; AAA6421.1; -
DR HSSP: P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 530 AA; 59564 MW; A78EC9C0D6268E58 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLGV 33
Db 288 LSEIKGIVHRLGV 302
|||||

RESULT 7
P90331 PRELIMINARY; PRT; 550 AA.
AC P90331;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGAHATA;
RA Sheng J., Watanabe M., Ueda S.;
RT "Selection of a neurotropic variant of measles virus."
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGAHATA;
DT 01-MAY-2000 (TREMblrel. 13, Created)

RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced
RT fusogenicity of measles virus."
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGAHATA;
RA Ning X., Ayala M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: D63926; BAA0958.1; -
DR EMBL: AF179431; AAF02696.1; -
DR HSSP: P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59530 MW; 97C991C7E2169839 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLGV 33
Db 288 LSEIKGIVHRLGV 302
|||||

RESULT 8
OY 091094 PRELIMINARY; PRT; 550 AA.
AC 091094;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-9301V;
RX MEDLINE-98440529; PubMed-9765410;
RA Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shiota T., Sakai Y.,
RA Asakawa M., Nagai Y.;
RT "Measles virus attenuation associated with transcriptional impediment
RT and a few amino acid changes in the polymerase and accessory
RT proteins."
RL J. Virol. 72:8690-8696(1998).
DR EMBL: AB012949; BAA33877.1; -
DR EMBL: AB012948; BAA33871.1; -
DR HSSP: P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59512 MW; 7AA4F1D117197BF9 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLGV 33
Db 288 LSEIKGIVHRLGV 302
|||||

RESULT 9
OY 090EX1 PRELIMINARY; PRT; 550 AA.
AC 090EX1;
DT 01-MAY-2000 (TREMblrel. 13, Created)

```

RESULT 2
ID 004243 PRELIMINARY; PRT; 534 AA.
AC 004243;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,
RA Billeter M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RN Cell 55:255-265(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billeter M.A.;
RA Virology 0:0-0(0).
DR EMBL; X16568; CAA34581.1; -.
DR EMBL; X16568; CAA34582.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 534 AA; 57899 MW; 637245E23B5BE044 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIYHRLGCV 33
Db 291 LSEIKGVIYHRLGCV 305

RESULT 3
ID 004242 PRELIMINARY; PRT; 537 AA.
AC 004242;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,
RA Billeter M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RN Cell 55:255-265(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billeter M.A.;
RA Virology 0:0-0(0).
DR EMBL; X16567; CAA34574.1; -.
DR EMBL; X16567; CAA34575.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.

```

```

SQ SEQUENCE 537 AA; 58275 MW; D0A60AC6D979E06 CRC64;

Query Match 29.0%; Score 72; DB 12; Length 537;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIYHRLGCV 33
Db 291 LSEIKGVIYHRLGCV 305

RESULT 4
ID 09PXA4 PRELIMINARY; PRT; 545 AA.
AC 09PXA4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=OSA-3;
RC Ning X., Ayala M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179440; AAF02705.1; -.
DR EMBL; AF179439; AAF02704.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 545 AA; 58907 MW; 0234C28AE193E77D CRC64;

Query Match 29.0%; Score 72; DB 12; Length 545;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIYHRLGCV 33
Db 288 LSEIKGVIYHRLGCV 302

RESULT 5
ID 091HA5 PRELIMINARY; PRT; 546 AA.
AC 091HA5;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Rinderpest virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11241;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K;
RX MEDLINE=21014265; PubMed=11186456;
RA Alanot P.K., Smirnov A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA Gusev A.A.;
RT "Primary structure of the F-gene from Rinderpest virus strain K.";
RL Mol. Gen. Microbiol. Virusol. 4:29-33(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K;

```



```

RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE-83126573; PubMed-6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
RT biosynthesized in the human placenta.";
RL Biochem Biophys. Res. Commun. 109:1061-1071(1992).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC Lutepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm
CC (Serono).
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X01059; CAA25526.1; -
DR EMBL: M12578; AAA35916.1; -
DR EMBL: X15215; CAA33285.1; -
DR PIR: A01410; RHHUG.
DR PIR: A26173; A26173.
DR PIR: S05308; S05308.
DR MIM: 152760; -
DR InterPro: IPR002012; GnRH.
DR InterPro: IPR004079; Gonadoliberin1.
DR Pfam: PF00446; GnRH; 1.
DR PRINTS: PR01541; GONADOLIBERN1.
DR PROSITE: PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal.
FT SIGNAL 1 23 PROGONADOLIBERIN 1.
FT CHAIN 24 92 GONADOLIBERIN 1.
FT PEPTIDE 24 33 GONADOLIBERIN 1.
FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE 1.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 24 24 PYROLIDONE CARBOXYLIC ACID.
FT RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT CONFLICT 16 16 W -> S (IN REF. 3).
SQ SEQUENCE 92 AA; 10380 MM; 30A72221B076FA79 CRC64;
Query Match 25.8%; Score 64; DB 1; Length 92;
Best Local Similarity 40.0%; Pred. No. 0.17;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
OY 2 HWSGLRPPSSGSLKLSEIKGIY 26
Db 25 HWSGLRPPGKGRKAEMLIDSPQETV 49
RESULT 14
VGLF_MEAST STANDARD; PRT; 529 AA.
AC P26031; Q83298;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain IP-3-Ca) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11237;
```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92263801; PubMed-1585658;
RA Schmid A., Spielhofer P., Cattaneo R., Baczko K., Ter Meulen V.,
RA Billeter M.A.;
RT "Subacute sclerosing panencephalitis is typically characterized by
RT alterations in the fusion protein cytoplasmic domain of the
RT persisting measles virus.";
RL Virology 188:910-915(1992).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16566; CAA34567.1; -
DR EMBL: X16566; CAA34568.1; ALT_INIT.
DR HSSP: P04849; ISVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 26 BY SIMILARITY.
FT CHAIN 27 529 FUSION GLYCOPROTEIN F0.
FT CHAIN 27 115 PROTEIN F2.
FT CHAIN 116 529 PROTEIN F1.
FT TRANSMEM 116 139 POTENTIAL.
FT DOMAIN 140 497 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 498 518 POTENTIAL.
FT DOMAIN 519 529 CYTOPLASMIC (POTENTIAL).
FT DISULFID 71 198 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 32 32 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 529 AA; 57331 MM; AE987BC9F07E9A9 CRC64;
Query Match 25.8%; Score 64; DB 1; Length 529;
Best Local Similarity 93.3%; Pred. No. 1.1;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 19 LSEIKGIYVHRLGCV 33
Db 291 LSEIKGIYVHRLGCV 305
RESULT 15
GONI_HAPBU STANDARD; PRT; 94 AA.
AC P51918; O93387;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin 1 precursor (gonadotrophin-releasing hormone 1) (GnRH-1)
DE (LH-RH 1) (Luliberin 1).
GN GnRH1.
OS Haplochromis burtoni (Burton's mouthbrooder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Astatotilapia.
OX NCBI_TaxID=8153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95396797; PubMed-7667296;
RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D10371; BAA01206.1; -.
CC PIR: J01368; VGNZPD.
CC PIR: A48346; A48346.
CC HSSP: P04849; 1SVF.
CC InterPro: IPR000776; Fusion_gly.
CC Pfam: PF00523; Fusion_gly.1.
CC GlycoProtein: Fusion protein; Transmembrane; Envelope protein; Signal.
CC SIGNAL 1 2
CC CHAIN ? 631 FUSION GLYCOPROTEIN F0.
CC CHAIN ? 188 F2 PROTEIN.
CC CHAIN 194 631 F1 PROTEIN.
CC DISULFID 149 276 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
CC TRANSMEM 89 106 POTENTIAL.
CC TRANSMEM 194 212 POTENTIAL.
CC TRANSMEM 575 595 POTENTIAL.
CC CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CONFLICT 63 63 I -> V (IN REF. 2).
CC SEQUENCE 631 AA; 68873 MW; D1FC87CDD426E9B8 CRC64;
SO
Query Match 26.2%; Score 65; DB 1; Length 631;
Best Local Similarity 68.4%; Pred. No. 1;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 15 SLKLSEIKGYIVHLEGV 33
DB 365 SYPTLSEKGVVHLEAV 383

```

```

CC -----
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S75918; AAB33096.1; -.
CC InterPro: IPR002012; GNRH.
CC Pfam: PF00446; GNRH.1.
CC PROSITE: PS00473; GNRH.1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal 1 1
CC NON_TER 1 1 BY SIMILARITY.
CC SIGNAL <1 5 PROGONADOLIBERIN I.
CC CHAIN 6 6 GNRADOLIBERIN I.
CC PEPTIDE 6 15 GNRH-ASSOCIATED PEPTIDE I.
CC ACT_SITE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
CC ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC ACTIVITY (BY SIMILARITY).
CC MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
CC MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
CC MOD_RES 15 15 SIMILARITY).
CC NON_TER 67 67
CC SEQUENCE 67 AA; 7573 MW; 505394DA261A3F2 CRC64;
SO
Query Match 25.8%; Score 64; DB 1; Length 67;
Best Local Similarity 40.0%; Pred. No. 0.12;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
OY 2 HWSYGLRPGSSPSLKLSEIKGYI 26
DB 7 HWSYGLRPGKRDALNLDSPFQETV 31

```

```

RESULT 12
CONL_MACMU STANDARD; PRT; 67 AA.
ID CONL_MACMU STANDARD; PRT; 67 AA.
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progondoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I]
DE (Fragment).
DE GNRH1 OR GNRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RT Neuroendocrinology 60:346-359(1994).
RT
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

RESULT 13
CONL_HUMAN STANDARD; PRT; 92 AA.
ID CONL_HUMAN STANDARD; PRT; 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Progondoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated
DE peptide I].
DE GNRH1 OR GNRH OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=8936682; PubMed=2671939;
RX Haylick J.S., Adelman J.P., Seeburg P.H.;
RA Haylick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RT hormone gene.";
RT Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=86094338; PubMed=2867548;
RX Adelman J.P., Mason A.J., Haylick J.S., Seeburg P.H.;
RA Adelman J.P., Mason A.J., Haylick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RT Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=85012739; PubMed=6090951;
RX Seeburg P.H., Adelman J.P.;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone.";
RT Nature 311:666-668(1984).

```

```

DR EMBL: X65509; CAA46481.1; -
DR PIR: JS0321; VGNZCD.
DR HSSP: P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 662 FUSION GLYCOPROTEIN F0.
FT CHAIN 1 662 PROTEIN F2.
FT CHAIN 2 224 PROTEIN F1.
FT TRANSMEM 606 629 POTENTIAL.
FT DISULFID 180 307 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 3 3 R -> K (IN REF. 2).
FT CONFLICT 140 140 D -> N (IN REF. 2).
FT CONFLICT 152 152 N -> S (IN REF. 2).
FT CONFLICT 171 171 I -> M (IN REF. 2).
FT CONFLICT 174 174 A -> V (IN REF. 2).
FT CONFLICT 662 662 L -> H (IN REF. 2).
SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;

Query Match 26.5%; Score 66; DB 1; Length 662;
Best Local Similarity 73.7%; Pred. No. 0.81;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 15 SLKLSKGVYHRLGCV 33
DB 396 SYPTLSEKGVYHRLGCV 414

RESULT 10
VGLE_RINDK STANDARD; PRT; 546 AA.
AC P12574;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain Kabete O) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=11242;
RN RP
RX MEDLINE=88322864; PubMed=3413983;
RA Hsu D., Yamanaka M., Miller J., Dale B., Grubman M., Yima T.;
RT "Cloning of the fusion gene of rinderpest virus: comparative sequence
RT analysis with other morbilliviruses.";
RL Virology 166:149-153(1988).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M21514; AAA47400.1; -
CC PIR: A31051; VGNZMK.
CC HSSP: P04849; 1SVF.
CC InterPro: IPR000776; Fusion_gly.

```

```

DR Pfam: PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 546 F2 PROTEIN.
FT CHAIN 20 108 F1 PROTEIN.
FT CHAIN 109 546
FT DOMAIN 104 108 ARG-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 ARG/LYS-RICH (BASIC).
FT DOMAIN 514 517 POTENTIAL.
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58662 MW; 476D74DC18BCFCF CRC64;

Query Match 26.2%; Score 65; DB 1; Length 546;
Best Local Similarity 86.7%; Pred. No. 0.88;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 19 LSEIKGVYHRLGCV 33
DB 284 LSEIKGVYHRLGCV 298

RESULT 11
VGLE_PHODV STANDARD; PRT; 631 AA.
AC P28886;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Phocine distemper virus (PDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=11240;
RN RP
RX MEDLINE=92113538; PubMed=1765768;
RA Koevaanes J., Blixenkron-Moeller M., Sharma B., Oerfell C.,
RA Norby E.;
RT "The nucleotide sequence and deduced amino acid composition of the
RT haemagglutinin and fusion proteins of the morbillivirus phocid
RT distemper virus.";
RL J. Gen. Virol. 72:2959-2966(1991).
RN RN
RX MEDLINE=92398437; PubMed=1524494;
RA Curran M.D., Lu Y.J., Rima B.K.;
RT "The fusion protein gene of phocine distemper virus: nucleotide and
RT deduced amino acid sequences and a comparison of morbillivirus fusion
RT proteins.";
RL Arch. Virol. 126:159-169(1992).
RN RN
RX MEDLINE=91089508; PubMed=2264246;
RA Curran M.D., Ioan D.O., Rima B.K., Kennedy S.;
RT "Nucleotide sequence analysis of phocine distemper virus reveals its
RT distinctness from canine distemper virus.";
RL Vet. Rec. 127:430-431(1990).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.

```


RA Biller M.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14915; AAA64423.1; -;
DR EMBL; X05597; CAA29090.1; ALT_INIT.
DR EMBL; K01711; AAA75498.1; ALT_INIT.
DR EMBL; K01711; AAA75499.1; -;
DR EMBL; U03657; AAA56647.1; ALT_INIT.
DR EMBL; U03659; AAA56649.1; ALT_INIT.
DR EMBL; U03670; AAA56660.1; ALT_INIT.
DR EMBL; U08416; AAA50550.1; ALT_INIT.
DR EMBL; Z66517; CAA91367.1; ALT_INIT.
DR EMBL; Z66517; CAA91368.1; -;
DR PIR; A26962; VGNZMV.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 550 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT DISULFID 58 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA; 59532 MW; 7AAAF1CA82169093 CRC64;
Query Match 29.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LSEIKGVIHRLEGV 33
ID LSEIKGVIHRLEGV 33
DB 288 LSEIKGVIHRLEGV 302
RESULT 6
ID VGLF_RINDR STANDARD; PRT; 546 AA.
AC P41356;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN GNRH1
OS Rinderpest virus (strain RBOK (RDV)).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCB1_TaxID=36409;
RN [1]
RP MEDLINE=95088609; PubMed=7996154;
RA Evans S.A.; Baron M.D.; Chamberlain R.W.; Goatley L.; Barrett T.;

RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus."
RL J. Gen. Virol. 75:3611-3617(1994).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z30700; CAA83186.1; -;
DR EMBL; Z30697; CAA83181.1; -;
DR PIR; S47305; S47305.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 ARG/LYS-RICH (BASIC).
FT DOMAIN 514 517 POTENTIAL.
FT DISULFID 54 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58705 MW; ED3DF8AFEDBCB95 CRC64;
Query Match 28.6%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.15;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 19 LSEIKGVIHRLEGV 33
ID LSEIKGVIHRLEGV 33
DB 284 LSEIKGVIHRLEGV 298
RESULT 7
ID GON1_PAGMA STANDARD; PRT; 95 AA.
AC P70074;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I precursor [Gonadotropin-releasing hormone I] (GNRH-I)
DE (LH-RH I) (Lutiberin I).
GN GNRH1
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Sparidae; Pagrus.
OX NCB1_TaxID=143350;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Brain;
RA Okuzawa K.; Granneman J.; Bogerd J.; Goos H.; Zohar Y.; Kagawa H.;
RA Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.

FT MOD.RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
FT MOD.RES 35 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
SQ SEQUENCE 95 AA; 10753 MW; 49313FD6FDB87DA CRC64;
Query Match 29.2%; Score 72.5; DB 1; Length 95;
Best Local Similarity 45.7%; Pred. No. 0.015;
Matches 16; Conservative 3; Mismatches 15; Indels 1; Gaps 1;
Db 2 HWSGLRPGSSGSLKLSEIKGVIYHRLGEGV 36
27 HWSGLSPGSK-RDLDSLSTGLNITFRPHVDSF 60
RESULT 2
VGLE_MEASY STANDARD; PRT; 534 AA.
ID VGLE_MEASY
AC P26032;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
Fusion glycoprotein F1].
GN F.
OS Measles virus (strain Yamagata-1) (Subacute sclerose panencephalitis
virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=11239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90385702; PubMed=1698327;
RA Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;
RT "Molecular analysis of structural protein genes of the Yamagata-1
strain of defective subacute sclerosing panencephalitis virus. IV.
RT Nucleotide sequence of the fusion gene.";
RT Virus Genes 4:173-181(1990).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10548; BAA01405.1; -;
DR PIR; J00274; J00274.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly.1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 534 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 534 PROTEIN F2.
FT CHAIN 113 534 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 534 CYTOPLASMIC (POTENTIAL).
FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 534 AA; 57963 MW; F5B21757E643844D CRC64;
Query Match 29.0%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 19 LSEIKGVIYHRLGEGV 33
Db 288 LSEIKGVIYHRLGEGV 302
RESULT 3
VGLE_RINDB STANDARD; PRT; 546 AA.
ID VGLE_RINDB
AC P41360;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain RBT1) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OX NCBI_TaxID=39007;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088609; PubMed=7996154;
RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus.";
RT J. Gen. Virol. 75:3611-3617(1994).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z31656; CAA83482.1; -;
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly.1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 546 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 O-LINKED (POTENTIAL).
SQ SEQUENCE 546 AA; 58418 MW; 38B539B89344F401 CRC64;
Query Match 29.0%; Score 72; DB 1; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 19 LSEIKGVIYHRLGEGV 33
Db 284 LSEIKGVIYHRLGEGV 298
RESULT 4

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 6.12607 Seconds
(without alignments)
297,061 Million cell updates/sec

Title: US-09-848-834a-17
Perfect score: 248
Sequence: 1 XHMSYGLRPGSSGSPSLKLS.....HRLGVGPGSLHMSYGLRFX 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	# Match	Length	DB	ID	Description
1	72.5	29.2	95	1	GONI_SPAU	P51919 sparus aura
2	72	29.0	534	1	VGLE_MEASY	P26032 measles vir
3	72	29.0	546	1	VGLE_RINDB	P41360 rinderpest
4	72	29.0	550	1	VGLE_MEASA	P35973 measles vir
5	72	29.0	550	1	VGLE_MEASA	P08300 measles vir
6	71	28.6	546	1	VGLE_MEASE	P41356 rinderpest
7	68.5	27.6	95	1	GONI_RINDK	P70074 pagrus majo
8	66	26.6	546	1	VGLE_RINDL	P10864 rinderpest
9	66	26.6	662	1	VGLE_CDOVO	P12569 canine dist
10	65	26.2	546	1	VGLE_RINDK	P12574 rinderpest
11	65	26.2	631	1	VGLE_PHODY	P28886 phocine dis
12	64	25.8	67	1	GONI_MACMU	P55247 macaca mula
13	64	25.8	92	1	GONI_HUMAN	P01148 homo sapien
14	64	25.8	529	1	VGLE_MEASI	P36031 measles vir
15	63.5	25.6	94	1	GONI_HABBU	P51918 haplochromi
16	62.5	25.2	89	1	GONI_XENLA	P45656 xenopus lae
17	61.5	24.8	61	1	GONI_SHEEP	P28588 ovis aries
18	60.5	24.4	74	1	GONI_ONCMY	P55246 oncorhynch
19	60.5	24.4	82	1	GONI_SALTR	P5653 salmo trutt
20	60.5	24.4	880	1	TYO3_MOUSE	P55144 mus musculu
21	60	24.2	92	1	GONI_TURGB	P05335 tupaya glis
22	59.5	24.0	90	1	GONI_MOUSE	P13562 mus musculu
23	59.5	24.0	91	1	GONI_PIG	P49921 sus scrofa
24	59	23.8	95	1	GONI_MORSA	P073812 morone saxa
25	58	23.4	63	1	GONI_MESAU	P09163 mesocricetu
26	58	23.4	92	1	GONI_RAT	P07490 rattus norv
27	58	23.4	99	1	GONI_DICLA	P09140 dicentrarch
28	57.5	23.2	880	1	TYO3_RAT	P55146 rattus norv
29	56.5	22.8	74	1	GONI_ONCTS	P52097 oncorhynch
30	56.5	22.8	82	1	GONI_SALSA	P56629 salmo salar
31	55.5	22.4	233	1	MTRD_MERTM	P07230 methanobact
32	55.5	22.4	233	1	MTRD_MERTM	P07230 methanobact
33	55	22.2	92	1	GONI_CHICK	P37042 gallus gall

34	55	22.2	379	1	PURK_BACSU	P12045 bacillus su
35	54.5	22.0	249	1	PRA_MYCLE	P41484 mycobacteri
36	54	21.8	10	1	GONI_ALIMI	P37041 alligator m
37	53	21.4	497	1	DLHD_MANSE	P018480 manduca sex
38	53	21.4	508	1	VGIG_IHNY	P07923 infectious
39	52.5	21.2	217	1	RS3_BACST	P23309 bacillus st
40	52.5	21.2	481	1	DOKI_HUMAN	P09704 homo sapien
41	52	21.0	213	1	PPCT_BOVIN	P02720 bos taurus
42	52	21.0	520	1	YRV2_CAEBL	Q27514 caenorhabdi
43	52	21.0	1331	1	APCE_ANASP	P80559 anabaena sp
44	51.5	20.8	90	1	GON8_RANDY	Q91au2 rana dybows
45	51.5	20.8	207	1	TER2_ECOLI	P04483 escherichia

ALIGNMENTS

RESULT 1	GONI_SPAU	STANDARD:	PRT:	95 AA.
AC	P51919;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)			
DE	(LH-RH I) (Lutaliberin I) (SBGNRH).			
GN	GNRH1.			
OS	Sparus aurata (Gillthead sea bream).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorphi; Acanthopterygii; Perciformes; Percoidae;			
OC	Sparidae; Sparus.			
OK	NCBI_TaxID=8175;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=95268499; PubMed=7749463;			
RA	Gothliff Y., Elizur A., Chow M., Chen T.T., Zohar Y.,			
RT	"Molecular cloning and characterization of a novel gonadotropin-			
RT	releasing hormone from the gillhead seabream (Sparus aurata).";			
RT	Mol. Mar. Biol. Biotechnol. 4:27-35(1995).			
RL	[2]			
RP	SEQUENCE OF 26-35.			
RC	TISSUE=Brain;			
RX	MEDLINE=95083645; PubMed=7991588;			
RA	Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,			
RA	Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;			
RT	"Three forms of gonadotropin-releasing hormone characterized from			
RT	brains of one species.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).			
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.			
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: U30320; AAA75469.1; -			
DR	InterPro: IPR002012; GNRH.			
DR	Pfam: PF00446; GNRH. 1.			
DR	PROSITE, PS00473; GNRH. 1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;			
FT	SIGNAL; Multigene family.			
FT	CHAIN 1 25			
FT	PEPTIDE 26 35			
FT	PEPTIDE 39 95			
FT	PROGONADOLIBERIN I.			
FT	GONADOLIBERIN I.			
FT	PEPTIDE 39 95			
FT	GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).			

C:Superfamily: parainfluenza virus cell fusion protein
 C:Keywords: glycoprotein, membrane fusion, transmembrane protein
 E:1-188/0/Product: cell fusion glycoprotein F2 #status predicted <FP2>
 E:89-106/Product: transmembrane #status predicted <TM1>
 E:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
 F:194-219/Domain: transmembrane #status predicted <TM2>
 F:575-595/Domain: transmembrane #status predicted <TM3>
 F:110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	26.2%;	Score 65;	DB 1;	Length 631;
Best Local Similarity	68.4%;	Pred. No. 4.5;		
Matches 13;	Conservative 2;	Mismatches 4;	Indels 0;	Gaps 0;

QY 15 SLKLSEIKGVVHRLEGV 33
| |||:||||| |
Db 365 SYPTLSEVKGVVVHRL EAV 383

Search completed: October 10, 2002, 16:12:13
Job time : 12.0235 secs

VG2RL
 cell fusion glycoprotein precursor - rinderpest virus (strain L)
 N:Contains: fusion glycoprotein F1; fusion glycoprotein F2
 C:Species: rinderpest virus
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
 C:Accession: A28921
 R:Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.
 A:Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of the
 A:Reference number: A28921; MUID:88219541
 A:Accession: A28921
 A:Molecule type: mRNA
 A:Residues: 1-546 <TSU>
 A:Cross-references: GB:M20870; NID:g333898; PIDN:AAA47399.1; PID:g333899
 C:Genetics:
 A:Gene: F
 C:Superfamily: paramfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>
 F:105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>
 F:109-133/Domain: transmembrane #status predicted <TN1>
 F:485-513/Domain: transmembrane #status predicted <TN2>
 F:25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 1; Length 546;
 Best Local Similarity 93.3%; Pred. No. 2.9;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 19 LSEIKGVHRLGCV 33
 |||||
 Db 284 LSEIKGVHRLGCV 298

RESULT 12
 VG2CD
 cell fusion glycoprotein precursor - canine distemper virus
 N:Contains: fusion protein F1; fusion protein F2
 C:Species: canine distemper virus
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
 C:Accession: J50321
 R:Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.
 A:Title: The nucleotide sequence of the gene encoding the F protein of canine distemper
 A:Reference number: J50321; MUID:88129050
 A:Accession: J50321
 A:Molecule type: mRNA
 A:Residues: 1-662 <BAR>
 A:Cross-references: GB:M21849; NID:g323241; PIDN:AAA42878.1; PID:g323242
 C:Genetics:
 A:Gene: F
 C:Superfamily: paramfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein
 F:1-135/Domain: signal sequence #status predicted <SIG>
 F:136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>
 F:225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>
 F:606-629/Domain: transmembrane #status predicted <MEM>
 F:62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 1; Length 662;
 Best Local Similarity 73.7%; Pred. No. 3.6;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 15 SLKLSEIKGVHRLGCV 33
 |||||
 Db 396 SYPTLSEKGVHRLGCV 414

RESULT 13
 S21382
 cell fusion protein - canine distemper virus
 C:Species: canine distemper virus
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999

C:Accession: S21382
 R:Wild, T.F.; Bernard, A.; Spehner, D.; Villevall, D.; Drilling, R.
 Submitted to the EMBL Data Library, April 1992
 A:Description: Vaccination of mice against canine distemper virus induced encephalitis
 A:Reference number: S21382
 A:Accession: S21382
 A:Status: preliminary
 A:Molecule type: genomic RNA
 A:Residues: 1-662 <ML>
 A:Cross-references: EMBL:X65509; NID:g58853; PIDN:CAA46481.1; PID:g58854
 C:Superfamily: paramfluenza virus cell fusion protein

Query Match 26.6%; Score 66; DB 2; Length 662;
 Best Local Similarity 73.7%; Pred. No. 3.6;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 15 SLKLSEIKGVHRLGCV 33
 |||||
 Db 396 SYPTLSEKGVHRLGCV 414

RESULT 14
 VG2PD
 cell fusion glycoprotein precursor - phocine distemper virus
 N:Contains: fusion protein F1; fusion protein F2
 C:Species: phocine distemper virus
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996
 C:Accession: JQ1368
 R:Koevmees, J.; Blaxekrone-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.
 J. Gen. Virol. 72, 2959-2966, 1991
 A:Title: The nucleotide sequence and deduced amino acid composition of the haemagglut
 A:Reference number: JQ1368; MUID:92113538
 A:Accession: JQ1368
 A:Molecule type: genomic RNA
 A:Residues: 1-631 <NOV>
 C:Genetics:
 A:Gene: F
 C:Superfamily: paramfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein
 F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
 F:89-106/Domain: transmembrane #status predicted <TM1>
 F:189-193/Region: cleavage processing #status predicted
 F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
 F:194-212/Domain: transmembrane #status predicted <TM2>
 F:575-595/Domain: transmembrane #status predicted <TM3>
 F:110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.2%; Score 65; DB 1; Length 631;
 Best Local Similarity 68.4%; Pred. No. 4.5;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 15 SLKLSEIKGVHRLGCV 33
 |||||
 Db 365 SYPTLSEKGVHRLGCV 383

RESULT 15
 A48346
 cell fusion glycoprotein precursor - phocine distemper virus (strain Uster/88)
 N:Contains: fusion protein F1; fusion protein F2
 C:Species: phocine distemper virus
 C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 25-Oct-1996
 C:Accession: A48346
 R:Curran, M.D.; Lu, Y.J.; Rima, B.K.
 Arch. Virol. 126, 159-169, 1992
 A:Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced a
 A:Reference number: A48346; MUID:92398437
 A:Accession: A48346
 A:Molecule type: mRNA
 A:Residues: 1-631 <CUR>
 A:Note: sequence extracted from NCBI backbone (NCBITN:113098, NCBIPI:113099)
 C:Genetics:
 A:Gene: F

C:Keywords: glycoprotein; membrane fusion; transmembrane protein
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-110/Product: cell fusion glycoprotein F2 #status predicted <FP2>
 F:111-553/Product: cell fusion glycoprotein F1 #status predicted <FP1>
 F:501-517/Domain: transmembrane #status predicted <TMN>
 F:32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.0%; Score 72; DB 1; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0.53;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVVHRLEGV 33
 |||:|||||
 Db 291 LSEIKGVVHRLEGV 305

RESULT 7
 VGNZRK
 cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)

N:Contains: fusion glycoprotein F1; fusion glycoprotein F2
 C:Species: rinderpest virus

C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 25-Oct-1996
 C:Accession: A31051

R:Hu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.
 Virology 166, 149-153, 1988

A:Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis
 A:Reference number: A31051; MUID:88322864

A:Molecule type: genomic RNA
 A:Residues: 1-546 <HSU>

C:Genetics:
 A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein

F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-108/Product: cell fusion glycoprotein F2 #status predicted <FP1>

F:109-546/Product: cell fusion glycoprotein F1 #status predicted <FP2>
 F:109-134/Domain: transmembrane #status predicted <TMN>

F:491-513/Domain: transmembrane #status predicted <TMN>
 F:23,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.6%; Score 71; DB 1; Length 546;
 Best Local Similarity 93.3%; Pred. No. 0.7;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVVHRLEGV 33
 |||:|||||
 Db 284 LSEIKGVVHRLEGV 298

RESULT 8
 S47305
 gene F protein - rinderpest virus

C:Species: rinderpest virus
 C:Date: 20-Oct-1994 #sequence_revision 08-Sep-1995 #text_change 20-Sep-1999
 C:Accession: S47305; S47301

R:Baron, M.D.; Barrett, T.
 Submitted to the EMBL Data Library, March 1994

A:Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30
 A:Reference number: S47283

A:Molecule type: mRNA
 A:Residues: 1-546 <BAR>

A:Cross-references: EMBL:Z30697; NID:9535396; PIDN:CAA83181.1; PID:9535401; EMBL:Z30700;
 C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: transmembrane protein

Query Match 28.6%; Score 71; DB 2; Length 546;
 Best Local Similarity 93.3%; Pred. No. 0.7;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVVHRLEGV 33
 |||:|||||

Db 284 LSEIKGVVHRLEGV 298

RESULT 9
 S47034

cell fusion protein precursor - porpoise morbillivirus
 N:Alternate names: F protein

C:Species: porpoise morbillivirus
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
 C:Accession: S47034

R:Bolt, G.; Gottschalk, E.; Blixenkrone-Moeller, M.; Wisnaupt, R.G.A.; Welsh, M.J.;
 submitted to the EMBL Data Library, July 1994

A:Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbi
 A:Reference number: S47034

A:Molecule type: mRNA
 A:Residues: 1-552 <BOI>

A:Cross-references: EMBL:X80757; NID:9520639; PIDN:CAA56731.1; PID:9520640
 A:Experimental source: isolate Ulster 88

A>Note: the source is designated as Cetacean morbillivirus
 C:Superfamily: parainfluenza virus cell fusion protein

F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-552/Product: fusion protein #status predicted <MAT>

Query Match 26.8%; Score 66.5; DB 2; Length 552;
 Best Local Similarity 61.5%; Pred. No. 2.5;
 Matches 16; Conservative 1; Mismatches 2; Indels 7; Gaps 1;

OY 19 LSEIKGVVHRLEGVPSLHMSYGL 44
 |||:|||||
 Db 290 LSEIKGVVHRLEAV-----SYNL 308

RESULT 10
 JQ2223
 cell fusion protein F0 precursor - phocine distemper virus

N:Contains: F1 and F2 chains
 C:Species: phocine distemper virus

C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 24-Nov-1999
 C:Accession: JQ2223

R:Visser, I.K.G.; van der Heijden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oerly
 J. Gen. Virol. 74, 1989-1994, 1993

A:Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites
 e virus entity.

A:Reference number: JQ2223; MUID:93389459
 A:Accession: JQ2223

A:Molecule type: mRNA
 A:Residues: 1-542 <VTS>

A:Cross-references: GB:L07075
 A>Note: the authors translated the codon ATC for residue 4 as Leu
 C:Comment: This fusion protein F0 is cleaved into F1 and F2 chains.

C:Genetics:
 A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein

F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-99/Product: fusion protein #status predicted <MAT>

F:105-542/Product: F2 chain #status predicted <F2C>
 F:105-135/Region: hydrophobic
 F:486-512/Domain: transmembrane #status predicted <TMN>

F:21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.6%; Score 66; DB 2; Length 542;
 Best Local Similarity 73.7%; Pred. No. 2.9;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 15 SKLISEIKGVVHRLEGV 33
 |||:|||||
 Db 276 SYPLSEIKGVVHRLEAV/294

RESULT 11


```
RESULT 3
JU0274
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain
M:contans: fusion glycoprotein F1: fusion glycoprotein F2
C:Species: subacute sclerosing panencephalitis virus, SSPEV
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
C:Accession: JU0274
R:Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.
Virus Genes 4, 173-181, 1990
A:Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe
A:Reference number: JU0274; MUID:90385702
A:Accession: JU0274
A:Molecule type: mRNA
A:Residues: 1-534 <COM>
A:Cross-references: EMBL:DJ0548; NID:g222256; PID:BA01405.1; PID:g222257
A>Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F:108-534/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F:498-514/Domain: transmembrane #status predicted <TM>
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          29.0%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVVHRLGCV 33
Db 288 LSEIKGVVHRLGCV 302

RESULT 4
S47300
gene F protein - rinderpest virus
C:Species: rinderpest virus
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: S47300; PQ0865
R:Evans, S.A.; Barron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A:Description: The complete nucleotide sequence of the fusion protein gene of the vaccin
A:Reference number: S47299
A:Accession: S47300
A:Molecule type: DNA
A:Residues: 1-546 <EVA>
A:Cross-references: EMBL:Z31656; NID:g535406; PID:CAA83482.1; PID:g535407
R:Chamberlain, R.W.; Mamway, H.M.; Hockley, E.; Shalla, M.S.; Goatley, L.; Knowles, N.J.
J. Gen. Virol. 74, 2775-2780, 1993
A:Title: Evidence for different lineages of rinderpest virus reflecting their geographic
A:Reference number: PQ0865; MUID:94103786
A:Accession: PQ0865
A:Molecule type: mRNA
A:Residues: 86-191 <CHA>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match          29.0%; Score 72; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVVHRLGCV 33
Db 284 LSEIKGVVHRLGCV 298

RESULT 5
```

```
E48556
cell fusion glycoprotein precursor - measles virus (strain Aik-C)
C:Species: measles virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C:Accession: E48556
R:Morl, T.; Sasaki, K.; Hashimoto, H.; Makino, S.
Virus Genes 7, 67-81, 1993
A:Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the Aik
A:Reference number: A48556; MUID:93227570
A:Accession: E48556
A:Molecule type: genomic RNA
A:Residues: 1-550 <MOR>
A:Cross-references: GB:S58435; NID:g299460; PID:AA026145.1; PID:g299465
A>Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBIIP:129272)
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F:108-550/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F:113-138/Region: hydrophobic
F:495-514/Domain: transmembrane #status predicted <TM>
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          29.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGVVHRLGCV 33
Db 288 LSEIKGVVHRLGCV 302

RESULT 6
VGNZMV
cell fusion glycoprotein precursor - measles virus
C:Species: measles virus
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000
C:Accession: A26962; A25616; PQ0380; PQ0384
R:Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.
J. Gen. Virol. 68, 1695-1703, 1987
A:Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and co
A:Reference number: A92794; MUID:87224816
A:Accession: A26962
A:Molecule type: mRNA
A:Residues: 1-553 <BRC>
A:Cross-references: GB:D00090; NID:g222061; PID:BA00056.1; PID:g222062
A:Experimental source: Strain Halle
R:Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini
Virology 155, 508-523, 1986
A:Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles v
A:Reference number: A94350; MUID:87071668
A:Accession: A25616
A:Molecule type: mRNA
A:Residues: 4-553 <RIC>
A:Cross-references: GB:M14915; NID:g331762; PID:AAA46423.1; PID:g331763
A:Experimental source: strain Edmonston
R:Schultz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compari
A:Reference number: PQ0374; MUID:92300360
A:Accession: PQ0380
A:Molecule type: genomic RNA
A:Residues: 272-553 <SCH1>
A:Experimental source: isolate CL
A:Accession: PQ0384
A:Molecule type: genomic RNA
A:Residues: 272-553 <SCH2>
A:Experimental source: isolate SE
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 10.9466 Seconds

(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-17

Sequence: 1 XHMSYGLRPGSSGSPSLKLS.....HRLGEGPPLHMSYGLRPPX 47

Scoring table: BLAST62

Gapop 10.0 , Gapext 0.5

Searched: 28338 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 28338

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	29.0	282	2	P00376 cell fusion glycop
2	72	29.0	282	2	P00388 cell fusion glycop
3	72	29.0	534	1	J00274 cell fusion glycop
4	72	29.0	546	2	S47300 gene F protein - r
5	72	29.0	550	1	E48556 cell fusion glycop
6	72	29.0	553	1	VGNZMV cell fusion glycop
7	71	28.6	546	1	VGNZBK cell fusion glycop
8	71	28.6	546	2	S47305 gene F protein - r
9	66.5	26.8	552	2	S47034 cell fusion protei
10	66	26.6	542	2	J02223 cell fusion glycop
11	66	26.6	546	1	VGNZRL cell fusion glycop
12	66	26.6	662	1	VGNZCD cell fusion glycop
13	66	26.6	662	1	S21382 cell fusion glycop
14	65	26.2	631	1	VGNZPD cell fusion protei
15	65	26.2	631	1	A48346 cell fusion glycop
16	64	25.8	67	2	I78541 gonadoliberin prec
17	64	25.8	92	1	RHHUG gonadoliberin prec
18	63.5	25.6	98	2	I50739 gonadoliberin-relea
19	62.5	25.2	89	2	I51423 gonadoliberin prec
20	60.5	24.4	82	2	I51365 gonadotropin-relea
21	60.5	24.4	555	2	B72486 probable hydanoin
22	60.5	24.4	880	2	B53743 protein-tyrosine k
23	60	24.2	546	2	S55386 cell fusion protei
24	59.5	24.0	90	1	RHMSG gonadoliberin prec
25	59.5	24.0	1041	2	C83548 gonadoliberin prec
26	58.5	23.6	502	2	T36389 probable transmem
27	58	23.4	10	1	RHPGG gonadoliberin - sh
28	58	23.4	10	1	RHSHG gonadoliberin - sh
29	58	23.4	92	1	RHRTG gonadoliberin prec

30	57.5	23.2	509	2	T45529 agaA protein limpo
31	57.5	23.2	880	1	UC4166 protein-tyrosine k
32	57	23.0	371	2	T49908 hypothetical prote
33	56.5	22.8	74	2	I51092 gonadotropin relea
34	56.5	22.8	82	2	I51355 gonadotropin relea
35	56.5	22.8	82	2	I51331 gonadotropin relea
36	56	22.6	233	2	E87362 hypothetical prote
37	56	22.6	636	2	S47299 gene F protein - r
38	55.5	22.4	233	2	H69021 tetrahydromethanop
39	55.5	22.4	233	2	S38366 gonadoliberin I pr
40	55	22.2	92	2	I50644 phosphoribosylamin
41	55	22.2	379	1	DCBSPK probable galactosi
42	55	22.2	400	2	A40104 adenhyde reductase
43	54.5	22.0	190	2	T37168 probable tetr-fami
44	54.5	22.0	190	2	T37168 36k antigen pra -
45	54.5	22.0	249	2	A41497

ALIGNMENTS

RESULT 1

P00376 cell fusion glycoprotein - measles virus (strain Tt) (fragment)

C:Species: measles virus

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999

C:Accession: P00376

R.Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compari

A:Reference number: P00374; MUID:92300360

A:Accession: P00376

A:Molecule type: genomic RNA

A:Residues: 1-282 <SCH>

C:Genetics:

A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion

Query Match 29.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLGV 33

DB 20 LSEIKGIVHRLGV 34

RESULT 2

P00388 cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)

C:Species: measles virus

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999

C:Accession: P00388

R.Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compari

A:Reference number: P00374; MUID:92300360

A:Accession: P00388

A:Molecule type: genomic RNA

A:Residues: 1-282 <SCH>

C:Genetics:

A:Gene: F

C:Superfamily: parainfluenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion

Query Match 29.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LSEIKGIVHRLGV 33

DB 20 LSEIKGIVHRLGV 34

Job time : 8.61539 secs

TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..17
OTHER INFORMATION: /note= "Malaria circumsporozoite"
OTHER INFORMATION: 382-398
PCT-US95-02121-96

Query Match 43.2%; Score 79; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 KIAMKAKASSVFNVNS 36
Db 1 KIAMKAKASSVFNVNS 17

RESULT 15
US-08-817-933A-7
; Sequence 7, Application US/08817933A
; Patent No. 5945104
; GENERAL INFORMATION:
; APPLICANT: STANWORTH, DENIS R.
; APPLICANT: LEWIN, IAN V.
; TITLE OF INVENTION: PEPTIDES FOR ANTI-ALLERGY TREATMENT
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5945104th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,933A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9422294.0
; FILING DATE: 04-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 179-23
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4005
; TELEFAX: (703) 816-4100
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-817-933A-7

Query Match 41.0%; Score 75; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAMKAKASSVFNV 33
Db 1 EKKIAMKAKASSVFNV 16

Search completed: October 10, 2002, 16:14:05

ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-27

Query Match 48.6%; Score 89; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 3,7e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKIKAKKASSVFNVNS 36
|||||
Db 3 EKIKAKKASSVFNVNS 21

RESULT 13
US-08-488-351A-27
Sequence 27, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-27

Query Match 48.6%; Score 89; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 3,7e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKIKAKKASSVFNVNS 36
|||||
Db 3 EKIKAKKASSVFNVNS 21

RESULT 14
PCT-US95-02121-96
Sequence 96, Application PC/TUS9502121
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
TITLE OF INVENTION: CTL IMMUNITY
NUMBER OF SEQUENCES: 153
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02121
FILING DATE: 16-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,484
FILING DATE: 16-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,682
FILING DATE: 29-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,568
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: unknown

TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-54

Query Match 48.6%; Score 89; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36
|||||
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 10

PCT-US95-02121-97
Sequence 97, Application PC/TUS9502121
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
TITLE OF INVENTION: CTL IMMUNITY
NUMBER OF SEQUENCES: 153
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02121
FILING DATE: 16-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,484
FILING DATE: 16-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,682
FILING DATE: 29-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,568
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..21
OTHER INFORMATION: /note="Malaria circumsporozoite
OTHER INFORMATION: 378-398"

Query Match 48.6%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 EKKIAKMEKASSVFNVNS 36
|||||
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 11

PCT-US95-13841-20
Sequence 20, Application PC/TUS9513841
GENERAL INFORMATION:

APPLICANT: United Biomedical Inc; Wallfield, Alan M.;
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Synthetic IgE Membrane Anchor
TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13841
FILING DATE: 25-OCT-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/328,519
FILING DATE: 25-OCT-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lin, Maria C. H.
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4117
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-13841-20

Query Match 48.6%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36
|||||
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 12

US-08-446-692-27
Sequence 27, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:

APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:

US-08-446-692-48
; Sequence 48, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; ZIP: 10154-0053
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-48

Query Match 48.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 EKRIAKMERKASSVFNVYNS 36
DB 3 EKRIAKMERKASSVFNVYNS 21

RESULT 8
US-08-488-351A-48
; Sequence 48, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; ZIP: 10154-0053
; COUNTRY: US
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-48

Query Match 48.6%; Score 89; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 EKRIAKMERKASSVFNVYNS 36
DB 3 EKRIAKMERKASSVFNVYNS 21

RESULT 9
US-09-100-409A-54
; Sequence 54, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE: 514
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800

```

;
; TELEFAX: 610-270-5090
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-932-929B-1
;
Query Match          49.7%; Score 91; DB 4; Length 424;
Best Local Similarity 55.3%; Pred. No. 3.5e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGSPSID-----EKRIAKMEKASSVFNVYNS 36
::|||: 1 ||||| ||||| ||||| |||||
DB 156 IKGSAKPKDELDYANDIEKKICKMEKSSVFNVYNS 193

RESULT 5
US-08-932-929B-3
; Sequence 3, Application US/08932929B
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and Hbsag
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,929B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,797
; FILING DATE: 04-DEC-1996
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1FWC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-932-929B-3
;
Query Match          49.7%; Score 91; DB 4; Length 424;
Best Local Similarity 55.3%; Pred. No. 3.5e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGSPSID-----EKRIAKMEKASSVFNVYNS 36
::|||: 1 ||||| ||||| ||||| |||||
DB 152 IKGSAKPKDELDYANDIEKKICKMEKSSVFNVYNS 189
```

```

;
; RESULT 6
; US-08-186-266-6
; Sequence 6, Application US/08186266
; Patent No. 5662907
; GENERAL INFORMATION:
; APPLICANT: KUBO, Ralph T.
; APPLICANT: GREY, Howard M.
; APPLICANT: SETTE, Alessandro
; APPLICANT: CELIS, Esteban
; TITLE OF INVENTION: INDUCTION OF ANTI-TUMOR CYTOTOXIC
; TITLE OF INVENTION: T LYMPHOCYTES IN HUMANS USING
; TITLE OF INVENTION: SYNTHETIC PEPTIDE EPITOPES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,266
; FILING DATE: 25-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/159,339
; FILING DATE: 29-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-50-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION:
; OTHER INFORMATION: /note="Plasmodium falciparum CS
; protein at positions 378-398."
;
; US-08-186-266-6
;
Query Match          48.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKRIAKMEKASSVFNVYNS 36
||| ||||| ||||| ||||| |||||
DB 3 EKRIAKMEKASSVFNVYNS 21

RESULT 7
```



```
RESULT 2
US-08-760-797A-1
; Sequence 1, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-760-797A-1
Query Match 49.7%; Score 91; DB 2; Length 423;
Best Local Similarity 55.3%; Pred. No. 3.5e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAMERKASVFNVNS 36
      :|||:| 1 ||||| ||||| |||||
Db 155 IKPGSANPKDELVDYNDIEKKICKMEKSCSVFNVNS 192

RESULT 3
US-08-760-797A-3
; Sequence 3, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-760-797A-3
Query Match 49.7%; Score 91; DB 2; Length 424;
Best Local Similarity 55.3%; Pred. No. 3.5e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAMERKASVFNVNS 36
      :|||:| 1 ||||| ||||| |||||
Db 152 IKPGSANPKDELVDYNDIEKKICKMEKSCSVFNVNS 189

RESULT 4
US-08-932-929B-1
; Sequence 1, Application US/08932929B
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,929B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,797
; FILING DATE: 04-DEC-1996
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1FWC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
```


KM IGE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 KW vaccine; allergy; antibody; constant heavy chain.
 XX
 OS Plasmodium falciparum.
 PN WO9526365-A1.
 XX
 PD 05-OCT-1995.
 XX
 PF 24-MAR-1995; 95WO-US03741.
 XX
 PR 25-OCT-1994; 94US-0328912.
 PR 28-MAR-1994; 94US-0218461.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 PI
 DR WPI; 1995-351297/45.
 XX
 PT Synthetic peptide-based immunogen contg. IGE CH4 peptide and helper
 PT T cell epitope - useful for eliciting antibody prodn. for allergy
 PT treatment
 XX
 PS Claim 3; Page 23; 87pp; English.
 XX
 CC AAR82571-91 are helper T cell epitopes which can be used in the
 CC preparation of a peptide immunogen that is useful in vaccines for
 CC treating allergic reactions. In the immunogen an IGE CH4 peptide
 CC is attached C-terminally to a series of amino acids including a
 CC helper T cell epitope. The immunogen may also opt. contain a fatty
 CC acid or fatty acid derivative, an invasion domain or alpha-NH2. The
 CC immunogen produces high titres of antibodies to the effector site
 CC in human IGE heavy chain (the CH4 domain peptide) which inhibit mast
 CC cell activation and reduce allergen-induced IGE prodn. The immunogens
 CC may be used in either a radially branching multimeric form or a
 CC linearly arranged monomeric form.
 XX
 SQ Sequence 21 AA;
 Query Match 48.6%; Score 89; DB 16; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 EKRIAKKERASSYFNVYNS 36
 ||||||||||||||||
 DB 3 EKRIAKKERASSYFNVYNS 21

Search completed: October 10, 2002, 16:05:14
 Job time : 19.9231 secs

RESULT 13
 AAM98951
 ID AAM98951 standard; peptide; 19 AA.
 XX
 AC AAM98951;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Vaccine related MHC ligand peptide SEQ ID NO:54.
 XX
 KW Glutamic acid; glutamine; vaccine; major histocompatibility complex;
 KM MHC; immunomodulator; antiallergic; endocrine; neutroprotectant;
 KM virucidal; bactericidal; antiparasitic; fungicidal; cytostatic;
 KM medicine; pharmaceutical; immune disorder; immune deficiency;
 KM autoimmunity; hypersensitivity; allergy; graft rejection; infection;
 KM hormonal disorder; central nervous system disease; cancer; melanoma;
 KM anti-melanoma vaccine; human immunodeficiency virus.
 XX
 OS Plasmodium malariae.
 XX
 PN WO200170772-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 22-MAR-2001: 2001WO-FR00872.
 XX
 PR 23-MAR-2000: 2000FR-0003711.
 XX
 PA (FABR) FABRE MEDICAMENT SA PIERRE.
 XX
 PI Klingner-Hamour C, Corvaia N, Beck A, Goetsch L;
 DR WPI: 2001-611470/70.
 XX
 PT Stabilized pharmaceutical containing N-terminal glutamic acid or
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
 PT with strong acid -
 XX
 PS Claim 9: Page 39; 149pp; French.
 XX
 CC The present invention describes a pharmaceutical compound (I) that
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue
 CC in the form of an addition salt with a strong, physiologically
 CC acceptable acid (II). Also described are: (a) a pharmaceutical
 CC composition containing at least one (I); (b) a vaccine containing at
 CC least one (I) where this is a major histocompatibility complex (MHC)
 CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
 CC antiallergic, neutroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful, in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
 CC allergy, graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
 CC or fungal infections; or (ii) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I), e.g. melanoma
 CC and human immunodeficiency virus infection. AAM98951 to AAM99592
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.
 CC
 SQ Sequence 19 AA;
 XX
 Query Match 48.6%; Score 89; DB 22; Length 19;
 Best Local Similarity 100.0%; Pred. No. 3.9e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 EKRIAKMERKASVFNVNS 36
 ||||||||||||||||
 DB 1 EKRIAKMERKASVFNVNS 19

RESULT 14
 AAP91504
 ID AAP91504 standard; peptide; 21 AA.
 XX
 AC AAP91504;
 XX
 DT 13-MAR-1992 (first entry)
 XX
 DE Sequence of modified Plasmodium circumsporozoite T-cell epitope CS.T3.
 XX
 KW Circumsporozoite peptide; T-cell epitope; immunogenic composition;
 KM vaccine.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 1..2 /note= "May be H-Asp-Ile, H-Ile, or H-"
 FT MISC-difference 19..21 /note= "May be Val-Asn-Ser-OH, Val-Asn-OH,
 FT MISC-difference 19..21 /note= "May be Val-Asn-Ser-OH, Val-Asn-OH,
 FT MISC-difference 19..21 /note= "May be Val-Asn-Ser-OH, Val-Asn-OH,
 XX
 PN EP343460-A.
 XX
 PD 29-NOV-1989.
 XX
 PF 12-MAY-1989; 89EP-0108618.
 XX
 PR 24-MAY-1988; 88GB-0012214.
 XX
 PA (HOFF) HOFFMANN-LA ROCHE AG.
 XX
 PI Sinigaia F;
 DR WPI: 1989-349561/48.
 XX
 PT Modified Plasmodium CS peptide - used as a universally recognised
 PT T-cell epitope in vaccines to elicit an immune response against
 PT pathogenic agents
 XX
 PS Claim 1: Page 16; 23pp; English.
 XX
 CC Peptide CS.T3 (residues 1-21) corresponds to residues 378-398 of the
 CC CS protein from P. falciparum but contains 2 Ala residues in place
 CC of the native protein's Cys residues at positions 384 and 389. Also
 CC claimed is AAP91504 (or modified forms, see FT) associated with an
 CC antigenic structure representing a B-cell epitope, pref. a multiple
 CC antigenic peptide, esp. multimers of the repeat sequences NAMP
 CC present in P. falciparum CS protein.
 CC
 SQ Sequence 21 AA;
 XX
 Query Match 48.6%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 EKRIAKMERKASVFNVNS 36
 ||||||||||||||||
 DB 3 EKRIAKMERKASVFNVNS 21
 XX
 RESULT 15
 AAR82586
 ID AAR82586 standard; peptide; 21 AA.
 XX
 AC AAR82586;
 XX
 DT 13-JUN-1996 (first entry)
 XX
 DE Plasmodium falciparum circumsporozoite helper T cell epitope, PF.
 XX

CC	response.
XX	
SQ	Sequence 424 AA:
OY	7 INPGSSGPSLD-----EKKRIAMKESASYENVVNS 36 :: :: DB 152 IKPGSANKPKDELVDYANDIERKICKMEKSSVFENVNS 189
RESULT 11	
AAG63663	
ID	AAG63663 standard; peptide; 33 AA.
XX	
AC	AAG63663;
XX	
DT	29-OCT-2001 (first entry)
XX	
DE	Peptide comprising conjugation sites for a pseudopeptide.
XX	
KW	Pseudopeptide; immunomodulation; adjuvant; antigen presenting cell; macrophage; dendritic cell; vaccine; autoimmune disease.
OS	Synthetic.
XX	
PN	WO200146127-A1.
PD	28-JUN-2001.
XX	
PF	22-DEC-1999; 99WO-IB02038.
XX	
PR	22-DEC-1999; 99WO-IB02038.
PA	(OMP-H-) OM-PHARMA.
PI	Bauer J, Martin OR, Rodriguez S;
DR	WPI, 2001-502469/55.
XX	
PT	New amphiphilic acylated pseudopeptides having a functionalized auxiliary spacer, useful as immunomodulators, e.g. as adjuvants in vaccines -
XX	
PS	Example 3; Page 61; 166pp; French.
XX	
CC	The specification describes N-Acylated pseudopeptides, which have a neutral or charged acidic group at one terminal and a functionalized auxiliary spacer at the other. The pseudopeptides show immunomodulatory and adjuvant action, based on activation of antigen presenting cells (e.g. macrophages or dendritic cells), induction of differentiation of dendritic cells, induction of cytokine production and induction of maturation of immunocompetent cell strains originating from hematopoietic and lymphoid organs. They reinforce humoral and cellular immunity. They can be grafted onto antigens (to modulate immune response) or onto drugs (to improve the therapeutic activity or targeting). The pseudopeptides are thus useful in human or veterinary medicine as immunizing or diagnostic agents. Typically, they are used as adjuvants together with (or covalently bonded to) antigens for vaccination against viral, parasitic/protozoal, microbial or fungal infections; incubated with blood cells ex vivo, to render the cells immunocompetent before reintroduction in vivo; or used in therapy of certain autoimmune diseases. The pseudopeptides are useful as carriers for antigens or other therapeutic agents due to their ability to form non-covalent bonds via the hydrophobic or hydrophilic auxiliary spacer. The present sequence represents a peptide, which has conjugation sites for the pseudopeptides of the invention.
XX	
SO	Sequence 33 AA:

49.7%; Score 91; DB 14; Length 424;
Best Local Similarity 55.3%; Pred. No. 7.9e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1.

48.9%; Score 89.5; DB 22; Length 33;

RESULT 12
 AAG63516
 ID AAG63516 standard; peptide: 33 AA.
 XX
 AC AAG63516;
 XX
 DT 15-OCT-2001 (first entry)
 XX
 DE A peptide which may be conjugated to pseudopeptides.
 XX
 KW Pseudopeptide; immunomodulation; adjuvant; antigen presenting cell;
 KW macrophage; dendritic cell; cytokine production; immunocompetent cell;
 KW autoimmune disease.
 XX
 OS Synthetic.
 XX
 PN WO200146126-A1.
 PN
 PD 28-JUN-2001.
 PD
 XX
 PF 21-DEC-2000; 2000WO-FR03650.
 PF
 XX
 PR 22-DEC-1999; 99WO-IB02038.
 PR
 XX
 PA (OMPH-) OM-PHARMA.
 PA
 XX
 PI Bauer J, Marlin OR, Rodriguez S;
 PI
 DR WPI: 2001-496651/54.
 DR
 XX
 PT New amphiphilic acylated pseudopeptides having a functionalized
 PT auxiliary spacer, useful as immunomodulators, e.g. as adjuvants in
 PT vaccines -
 PS
 PS Example 3.4; Page 89; 267pp; French.
 CC
 CC The specification describes N-Acylated pseudopeptides, which have
 CC a neutral or charged acidic group at one terminal and a functionalized
 CC auxiliary spacer at the other. The pseudopeptides show immunomodulatory
 CC and adjuvant action, based on activation of antigen presenting cells
 CC (e.g. macrophages or dendritic cells), induction of differentiation of
 CC dendritic cells, induction of cytokine production and induction of
 CC maturation of immunocompetent cell strains originating from hematopoietic
 CC and lymphoid organs. They reinforce humoral and cellular immunity. They
 CC can be grafted onto antigens (to modulate immune response) or onto
 CC pseudopeptides are thus useful in human or veterinary medicine as
 CC immunizing or diagnostic agents. Typically, the pseudopeptides are used
 CC as adjuvants together with (or covalently bonded to) antigens for
 CC vaccination against viral, parasitic/protozoal, microbial or fungal
 CC infections; incubated with blood cells *ex vivo*, to render the cells
 CC immunocompetent before reintroduction *in vivo*; or used in therapy of
 CC certain autoimmune diseases. The present sequence represents a
 CC peptide which may be conjugated to pseudopeptides of the invention.
 CC
 XX
 SQ Sequence 33 AA:
 Query Match 48.9%; Score 89.5; DB 22; Length 33;
 Best Local Similarity 76.9%; Pred. No. 6.4e-06;
 Matches 20; Conservative 3; Mismatches 2; Indels 1; Gaps 1.
 11 SSGPSLDEKKIAKMEKASSVFNVNVS 36
 : : : : :
 Db 9 NANPDI-EKKIAKMEKASSVFNVNVS 33

ID	AAR37796
XX	AAR37796 standard; Protein: 424 AA.
AC	
XX	AAR37796;
D7	27-SEP-1993 (first entry)
XX	
DE	RfS protein.
KW	RfS; expression cassette; hybrid protein; S. cerevisiae; TDH3; cloning; circumsporozoite protein; CSP; Plasmodium falciparum; strain 7G8; hepatitis B virus; HBV; adw serotype; pres2 protein; S protein.
OS	Synthetic.
XX	
FH	Key
FT	Region 1 Location/Qualifiers
FT	/note= "Derived from S. cerevisiae TDH3 gene sequence"
FT	Region 2..4 /note= "Cloning artefact"
FT	Protein 5..193 /note= "Represents amino acids 210-398 of the CSP of P. falciparum"
FT	Region 194..197 /note= "Carboxy terminal amino acids from HBV (adw serotype) pres2 protein"
FT	Protein 198..424 /note= "S protein of HBV (adw serotype)"
PX	M09J10152-A.
PD	27-MAY-1993.
PE	11-NOV-1992; 92WO-EP02591.
PR	16-NOV-1991; 91GB-0024390. 27-FEB-1992; 92US-0842694.
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
Cohen J,	De Wilde M;
WT:	1993-182494/22.
N-PSDB;	AAQ42566.
HbYrd	protein comprising Plasmodium circumporozoite protein and HBAg - useful as a vaccine for treating patients susceptible to Plasmodium infections
Disclosure;	Fig 5; 59pp; English.
This	sequence represents the RfS hybrid protein which is encoded by the RfS expression cassette. This hybrid consists of a methionine residue derived from S. cerevisiae TDH3 gene sequence, three amino acids, Met-Ala-Pro, derived from a nucleotide sequence created by the cloning procedure used to construct the hybrid gene, a stretch of 189 amino acids representing amino acids 210 to 398 of the circumsporozoite protein (CSP) of Plasmodium falciparum strain 7G8, an amino acid Arg created by the cloning procedure, four amino acids, Pro-Val-Thr-Asn, representing the four carboxy terminal residues of hepatitis B virus (HBV), adw serotype, pres2 protein, and a stretch of 226 amino acids specifying the S protein of HBV, adw serotype. This protein, and RfS* (see also AAR37797), may be combined with an adjuvant and used in a vaccine for preventing plasmodium infections. The vaccines produce a humoral response and also a cellular immune response.
Sequence	424 AA;
Query Match	↓ 49.7%; Score 91; DB 14; Length 424; Best Local Similarity 55.3%; Pred. No. 7.9e-05;
Matches	21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

[illegible]

CC linked to DNA encoding Region II-contg. flanking region. The Pro
 CC residue separating the Asp (at the C-terminal of the linker) from
 CC the Region I-contg. CS flanking region is an artifact of a filled-
 CC in BamHI site; the Gly separating the Region I and II-contg. CS
 CC flanking regions is an artifact of a synthetic FokI/TthIII I
 CC linker. The peptide can be used in a vaccine for protection
 CC against malaria.
 CC See also AAR12306-R12311 and AAR13175-R13178.
 CC
 SO Sequence 335 AA;
 OY Query Match 49.7%; Score 91; DB 12; Length 335;
 DB Best Local Similarity 55.3%; Pred. No. 6e-05;
 Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 OY 7 LRPSSGSPSLD-----EKKIAMEKASVFNVNS 36
 DB 284 IKPGSANKPKDELVDYNDIEKICKMEKCSVFNVNS 321
 RESULT 7
 AAP83144
 ID AAP83144 standard; protein; 411 AA.
 XX
 AC AAP83144;
 XX
 DT 20-NOV-1990 (first entry)
 XX
 DE Sequence encoded by the circumsporozoite (CS) gene from
 DE Plasmodium falciparum.
 XX
 KM Vaccine; antigen; immunogen; probe; hybridisation;
 KM Immunoassay; diagnosis.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Region 106..120
 FT /note="Region 1"
 FT Region 147..206
 FT /note="Repeat region, repeat unit = NNP"
 FT Region 123..146
 FT /note="Repeat region, repeat unit = NNPVNDP"
 FT Region 211..286
 FT /note="Repeat region, repeat unit = NNP"
 XX
 PI EP278940-A.
 XX
 PD 17-AUG-1988.
 XX
 PF 25-JAN-1988; 88EP-0870008.
 XX
 PR 30-JAN-1987; 87US-0009325.
 XX
 PA (SMIK) SMITH KLINE-RIT.
 XX
 PI Cabezon T, De Wilde M, Harford N;
 XX
 DR WPI: 1988-229751/33.
 DR N-PSDB; AAN81108.
 XX
 FT DNA encoding hepatitis B virus antigens and hybrids contg. them -
 FT used for expression in yeast to obtain vaccines and bivalent
 PT vaccines
 XX
 PS Example; Fig 3Aa-3Af; 101pp; English.
 XX
 CC Sequence of the CS gene (AAN81108) is from lambda-mpI. A recombinant
 CC DNA molecule is claimed, comprising functional DNA coding sequence fused,
 CC in phase, to a portion of the Pre-S2 region of a hepatitis B virus (HBV)
 CC Pre-S2-S protein coding sequence. The functional DNA coding sequence
 CC comprises the Pre-S2 coding sequence, Pre-S1 coding sequence or entire
 CC Pre-S1-Pre-S2 coding sequence, the CS protein coding sequence or

CC Plasmodium, or a HIV coding sequence such as
 CC an HIV envelope gene sequence, e.g. HIV C7 protein coding region,
 CC peptide 121 coding region, or HIV Dreesman peptide coding region.
 XX
 SO Sequence 411 AA;
 OY Query Match 49.7%; Score 91; DB 9; Length 411;
 DB Best Local Similarity 55.3%; Pred. No. 7.6e-05;
 Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 OY 7 LRPSSGSPSLD-----EKKIAMEKASVFNVNS 36
 DB 360 IKPGSANKPKDELVDYNDIEKICKMEKCSVFNVNS 397
 RESULT 8
 AAP60416
 ID AAP60416 standard; protein; 412 AA.
 XX
 AC AAP60416;
 XX
 DT 13-JUN-1991 (first entry)
 XX
 DE CS protein of malaria parasite.
 XX
 KM Sporozoite; vaccination.
 XX
 OS Plasmodium falciparum.
 XX
 FH Key Location/Qualifiers
 FT Region 124..127
 FT /label= Repeat unit
 XX
 XX EP166410-A.
 XX
 PD 02-JAN-1986.
 XX
 PF 24-JUN-1985; 85EP-0107794.
 XX
 PR 26-JUN-1984; 84US-0624564.
 XX
 PA (USDC) US SEC OF COMMERCE.
 PA (USGO) US GOVERNMENT.
 PA (USSA) US SEC OF THE ARMY.
 XX
 PI McCutchan TF, Dame JB, Williams JL, Schneider I;
 XX
 DR WPI: 1986-008635/02.
 DR N-PSDB; AAN60362.
 XX
 PT New immunologically active pure synthetic peptide(s) - used for
 PT protection against infection by malaria parasite.
 XX
 PS Disclosure; Fig 2; 49pp; English.
 XX
 CC The Plasmodium CS gene was used to isolate peptides capable of
 CC inducing an immune response to the parasite. Peptide antigens may
 CC be synthesised in pure form and used to generate an immune
 CC response in vaccination against malaria. The featured repeat
 CC units are claimed and must be present in copies of 2-1000.
 XX
 SO Sequence 412 AA;
 OY Query Match 49.7%; Score 91; DB 7; Length 412;
 DB Best Local Similarity 55.3%; Pred. No. 7.6e-05;
 Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
 OY 7 LRPSSGSPSLD-----EKKIAMEKASVFNVNS 36
 DB 361 IKPGSANKPKDELVDYNDIEKICKMEKCSVFNVNS 398
 RESULT 9


```

KM hybrid; influenza virus; non-structural protein 1, fusion.
XX
OS Plasmodium falciparum.
XX Influenza virus (A/PR/8/34).
XX
FH Key
FT Region
FT 1..81
FT /label= "N-terminal of NS1
FT /note= "Influenza virus nonstructural protein 1"
FT 82..97
FT /label= "immunodominant repeat region
FT /note= "four tetrapeptide repeat units"
FT 98..103
FT /label= "synthetic linker
FT 104
FT /label= "artifact
FT /note= "see comments"
FT 105..209
FT /label= "AAs 19-123 of CS protein
FT /note= "Region I contg. flanking region less
FT signal sequence"
FT 210
FT /label= "artifact
FT /note= "see comments"
FT 211..335
FT /label= "AAs 288-412 of CS protein
FT /note= "Region II flanking region"
FT
XX
XX EP432965-A.
XX
XX 19-JUN-1991.
XX
XX 06-DEC-1990; 90EP-0313257.
XX
XX 08-DEC-1989; 89US-0447746.
XX
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF THE ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX
PI Gross MS, Gordon DM, Hollingdale MR;
XX
XX WPI: 1991-179771/25.
XX
PT Polypeptide comprising immunogenic determinants from P falciparum
PT - for vaccine against malaria infection in humans.
XX
PS Example 4; Page 11; 18pp; English.
XX
XX The polypeptide is prepared by genetic engineering of genes encoding
XX the P. falciparum circumsporozoite (CS) protein [Dame et al.,
XX Science 225 : 593 (1984)], and the influenza virus non-structural
XX protein 1 (NS1). [Baez et al., Nucleic Acids Research, 8 : 5845
XX (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX (NS1-01) is linked to a synthetic sequence encoding four repeat
XX units from the immunodominant region, which in turn is linked via
XX a synthetic sequence to DNA encoding Region I contg. flanking
XX region less the 18 AA signal region. This is linked to DNA
XX encoding Region II-contg. flanking region. The pro residue sep-
XX arating the ASP (at the C-terminal of the linker) from the Region
XX I-contg. CS flanking region is an artifact of a filled-in BamHI
XX site; the Gly separating the Region I and II-contg. CS flanking
XX regions is an artifact of a synthetic POKI/THIRI I linker. The
XX CC peptide can be used in a vaccine for protection against malaria.
XX See also AAR12306-RI2311 and AAR13175-RI3179.
XX
SQ Sequence 335 AA;
XX
Query Match, 49.7%; score 91; DB 12; Length 335;
Best local Similarity 55.3%; Pred. No. 6e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
7 LRPGSSGPSLD-----EKKIANMEKASSYENVVNS 36

```

```

Db      284 IKPGSANKPKDELXDYENDIEKICKCKMEKCSVFNVVNS 321
          ::|||:: | ||||| ||| ||||| |||
RESULT 6
AA13179
ID AAR13179 standard; Protein; 335 AA.
XX
AC AAR13179;
AD
DT 29-AUG-1991 (first entry)
DE NSL_81(NWDP)4RLFAuth.
KW Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
RW hybrid; influenza virus; non-structural protein 1; fusion.
XX
OS Plasmodium falciparum.
OS Influenza virus (A/PR/8/34).
XX
FH Key
FH Region
FT Location/Qualifiers
FT 1..81
FT /label= "N-terminal of NS1
FT /note= "Influenza virus nonstructural protein 1"
FT 82..97
FT /label= "immunodominant repeat region
FT /note= "four variant tetrapeptide repeat units"
FT 98..103
FT /label= "synthetic linker
FT 104
FT /label= "artifact
FT /note= "see comments"
FT 105..209
FT /label= "Aas 19-123 of CS protein
FT /note= "Region 1 contg. flanking region less
FT signal sequence"
FT 210
FT /label= "artifact
FT /note= "see comments"
FT 211..335
FT /label= "Aas 288-412 of CS protein
FT /note= "Region II flanking region"
FN EP432965-A.
PN
PE 19-JUN-1991.
PD
PE 06-DEC-1990; 90EP-0313257.
XX
PR 08-DEC-1989; 89US-0447746.
PA (SMIK ) SWITKLINE BEECHAM.
PA (USSA ) US SEC OF THE ARMY.
PA (BIOM-) BIOMEDICAL RES INST.
XX
PI Gross MS, Gordon DM, Hollingdale MR;
DR WPI; 1991-179771/25.
XX
PT Polypeptide comprising immunogenic determinants from P falciparum
PS - For vaccine against malaria infection in humans.
XX
Example 5; Page 11; 18pp; English.
XX
The polypeptide is prepd. by genetic engineering of genes encoding
the P. falciparum circumsporozoite (CS) protein [Dame et al.,
Science 225 : 593 (1984)], and the influenza virus non-structural
protein 1 (NSI), [Baez et al., Nucleic Acids Research, 8 : 5845
(1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NSI
(NSI-81) is linked to a synthetic sequence encoding four repeat
units (the variant form) from the immunodominant region, which in
turn is linked via a synthetic sequence to DNA encoding Region I
contg. flanking region less the 18 AA signal region. This is

```

```

XX 06-DEC-1990: 90EP-0313257.
XX 08-DEC-1989: 89US-0447746.
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF THE ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX Gross MS, Gordon DM, Hollingdale MR;
XX WPI: 1991-179771/25.
XX Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX Example 2; Page 10; 18pp; English.
XX The polypeptide is prep'd. by genetic engineering of genes encoding
XX the P. falciparum circumsporozoite (CS) protein [Dame et al.,
XX Science 225 : 593 (1984)], and the Influenza virus non-structural
XX protein 1 (NS1), [Baer et al., Nucleic Acids Research, 8 : 5845
XX (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX (NS1_81) is linked via a synthetic sequence to DNA encoding Region
XX I contg. flanking regionless the 18 AA signal region, which in
XX turn is fused to DNA encoding Region II-contg. flanking region.
XX This CS fusion is designated RLIfauth. The Pro residue separating
XX the Asp (at the C-terminal of the linker) from RLIfauth is an arti-
XX fact of a filled in BamHI site; the Gly separating Region I and
XX Region II-contg. CS flanking regions is an artifact of a synthetic
XX FokI/NotI linker. The peptide can be used in a vaccine for
XX protection against malaria.
XX The complete nucleotide and AA sequences are given in EP-304720,
XX filed May 1, 1990.
XX See also AAR12306-R12311 and AAR13175-R13179.
XX SQ Sequence 319 AA;
XX
XX Query Match 49.7%; Score 91; DB 12; Length 319;
XX Best Local Similarity 55.3%; Pred. No. 5.6e-05;
XX Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
XX
XX QY 7 LRPGSSGSPSLD-----EKRIAKMEKASSVFNVNS 36
XX :|||: | ||| ||| ||| ||| |||
XX Db 268 IKPGSANKPKDELDYENDIEKICKMEKSSVFNVNS 305
XX
XX RESULT 4
XX AAR13177
XX ID AAR13177 standard; Protein: 327 AA.
XX
XX AC AAR13177;
XX
XX XX 29-AUG-1991 (first entry)
XX XX NS1_81-RLfauth + (NANP)2.
XX DE
XX XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX KW hybrid; Influenza virus; non-structural protein 1; fusion.
XX XX
XX OS Plasmodium falciparum.
XX Influenza virus (A/PR/8/34/).
XX
XX XX Key Location/Qualifiers
XX FH 1..81
XX FT Region /label= N-terminal of NS1
XX FT /note= "Influenza virus nonstructural protein 1"
XX FT Peptide 82..87
XX FT /label= synthetic linker
XX FT Region 88
XX FT /label= artifact
XX FT /note= "see comments"
XX FT 89..193
XX Region

```

```

FT FT /label= AAs 19-123 of CS protein
FT FT /note= "Region I contg. flanking region less
FT FT signal sequence"
FT FT Region
FT FT 194..201
FT FT /label= immunodominant repeat region
FT FT /note= "two tetrapeptide repeat units"
FT FT 202
FT FT /label= artifact
FT FT /note= "see comments"
FT FT Region
FT FT 203..327
FT FT /label= AAs 288-412 of CS protein
FT FT /note= "Region II flanking region"
XX
XX PN EP432965-A.
XX
XX XX 19-JUN-1991.
XX PD
XX XX 06-DEC-1990: 90EP-0313257.
XX PF
XX XX 08-DEC-1989: 89US-0447746.
XX PR
XX XX (SMIK ) SMITHKLINE BEECHAM.
XX PA (USSA ) US SEC OF THE ARMY.
XX PA (BIOM-) BIOMEDICAL RES INST.
XX
XX PI Gross MS, Gordon DM, Hollingdale MR;
XX
XX DR WPI: 1991-179771/25.
XX
XX PT Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX
XX PS Example 3; Page 10; 18pp; English.
XX
XX XX The polypeptide is prep'd. by genetic engineering of genes encoding
XX CC the P. falciparum circumsporozoite (CS) protein [Dame et al.,
XX CC Science 225 : 593 (1984)], and the Influenza virus non-structural
XX CC protein 1 (NS1), [Baer et al., Nucleic Acids Research, 8 : 5845
XX CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX CC (NS1_81) is linked via a synthetic sequence to DNA encoding Region
XX CC I contg. flanking region less the 18 AA signal region. This is
XX CC linked to a synthetic sequence encoding two repeat units from the
XX CC immunodominant region, which in turn is fused to DNA encoding
XX CC Region II-contg. flanking region. The Pro residue separating the
XX CC Asp (at the C-terminal of the linker) from the Region I-contg. CS
XX CC flanking region is an artifact of a filled in BamHI site; the Gly
XX CC separating the repeat units and the Region II-contg. CS flanking
XX CC region is an artifact of a synthetic FokI/NotI linker. The
XX CC peptide can be used in a vaccine for protection against malaria.
XX CC See also AAR12306-R12311 and AAR13175-R13179.
XX
XX SQ Sequence 327 AA;
XX
XX Query Match 49.7%; Score 91; DB 12; Length 327;
XX Best Local Similarity 55.3%; Pred. No. 5.8e-05;
XX Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;
XX
XX QY 7 LRPGSSGSPSLD-----EKRIAKMEKASSVFNVNS 36
XX :|||: | ||| ||| ||| ||| |||
XX Db 276 IKPGSANKPKDELDYENDIEKICKMEKSSVFNVNS 313
XX
XX RESULT 5
XX AAR13178
XX ID AAR13178 standard; Protein: 335 AA.
XX
XX AC AAR13178;
XX
XX XX 29-AUG-1991 (first entry)
XX XX NS1_81(NANP)4RLfauth.
XX DE
XX XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX KW

```

```

PN  EP432965-A.
XX
XX  19-JUN-1991.
PD
XX  06-DEC-1990; 90EP-0313257.
PF
XX  08-DEC-1989; 89US-0447746.
PR
XX  (SMIK ) SMITHKLINE BEECHAM.
PA  (USSA ) US SEC OF THE ARMY.
PA  (BIOM-) BIOMEDICAL RES INST.
XX
XX  Gross MS, Gordon DM, Hollingdale MR;
PI  WPI; 1991-179771/25.
DR
XX  Polypeptide comprising immunogenic determinants from P falciparum
PT  - for vaccine against malaria infection in humans.
XX
XX  Example 1; Page 7; 18pp; English.
XX
CC  The polypeptide is prepd. by genetic engineering of genes encoding
CC  the P. falciparum circumsporozoite (CS) protein (Dame et al.,
CC  Science 225 : 593 (1984)), and the influenza virus non-structural
CC  protein 1 (NS1), [Baer et al., Nucleic Acids Research, 8 : 5845
CC  (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
CC  (NS1_81) is linked via a synthetic sequence to DNA encoding Region
CC  I contg. flanking regionless the 18 AA signal region, which in
CC  turn is fused to DNA encoding Region II contg. flanking region
CC  less the first nine N-terminal AAs. This CS fusion is designated
CC  RfIdelta9. The pro residue separating the Asp (at the C-terminal
CC  of the linker) from RfIdelta9 is an artifact of a filled in BamHI
CC  site. The peptide can be used in a vaccine for protection against
CC  malaria.
CC  See also AAR12306-R12311 and AAR13176-R13179.
XX
SQ  Sequence 309 AA;

Query Match 49.7%; Score 91; DB 12; Length 309;
Best Local Similarity 55.3%; Pred. No. 5.4e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY  7 LRPGSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
    :::|||: | ||| ||| ||| ||| ||| |||
DB  258 IKPGSANKPKDELVDYENDIEKKICKMEKCSVFNVNS 295

RESULT 2
AAR07945
ID  AAR07945 standard; protein; 319 AA.
XX
XX  AAR07945;
AC
XX
XX  22-FEB-1991 (first entry)
DT
XX  NS181RLFAuth plasmid product.
DE
XX  Malaria; vaccine.
KM
XX  Plasmodium falciparum.
OS
XX
XX  Key
FH  Location/Qualifiers
FT  1..81
FT  /label= NS181 protein fragment
FT  /note= "from plasmid pMG-1"
FT  89..193
FT  /label= Fragment of circumsporozoite protein
FT  204..319
FT  /label= Fragment of circumsporozoite protein
XX
XX  EP398540-A.
XX
XX  22-NOV-1990.
PD

```

```

XX  01-MAY-1990; 90EP-0304720.
PF
XX  03-MAY-1989; 89US-0346863.
PR
XX  (SMIK ) SMITHKLINE BEECHAM.
PA
XX  Gross MS, Young JF;
PI  WPI; 1990-350299/47.
DR  N-PSDB; AAQ06580.
DR
XX  New polypeptide used in malaria vaccine - comprises immunogenic
PT  determinants from 1st and 2nd flanking regions of plasmodium
PT  surface protein and intermediate repeat domain
XX
XX  Example 2; Page 11-12; 24pp; English.
XX
XX  The product is useful in preparation of vaccines for treatment and
CC  prophylaxis of plasmodium sporozoite infection. It may be easily
CC  produced in large pure quantities from a transformed E. coli
CC  expression system.
XX
SQ  Sequence 319 AA;

Query Match 49.7%; Score 91; DB 11; Length 319;
Best Local Similarity 55.3%; Pred. No. 5.6e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY  7 LRPGSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
    :::|||: | ||| ||| ||| ||| ||| |||
DB  268 IKPGSANKPKDELVDYENDIEKKICKMEKCSVFNVNS 305

RESULT 3
AAR13176
ID  AAR13176 standard; protein; 319 AA.
XX
XX  AAR13176;
AC
XX
XX  29-AUG-1991 (first entry)
DT
XX  NS1_81-RLFAuth.
DE
XX  Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
KM  hybrid; influenza virus; non-structural protein 1; fusion.
KM
XX  Plasmodium falciparum.
OS  Influenza virus (A/PR/8/34).
XX
XX  Key
FH  Location/Qualifiers
FT  1..81
FT  /label= N-terminal of NS1
FT  /note= "Influenza virus nonstructural protein 1"
FT  82..87
FT  /label= synthetic linker
FT  88
FT  /label= artifact
FT  /note= "see comments"
FT  89..193
FT  /label= AAs 19-123 of CS protein
FT  /note= "Region I contg. flanking region less
FT  194
FT  /label= artifact
FT  /note= "see comments"
FT  195..319
FT  /label= AAs 288-412 of CS protein
FT  /note= "Region II flanking region"
XX
XX  EP432965-A.
XX
XX  19-JUN-1991.
PD

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 18.9231 Seconds
(without alignments)
211.311 Million cell updates/sec

Title:	US-09-848-834A-16
Perfect score:	183
Sequence:	1 XHMSYGLRPGSSGSPSLDEKKIAKMEKASSVFNNVNS 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

1:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT *
2:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT *
3:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT *
4:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT *
5:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT *
6:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT *
7:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT *
8:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT *
9:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT *
10:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT *
11:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT *
12:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT *
13:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT *
14:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT *
15:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT *
16:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT *
17:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT *
18:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT *
19:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT *
20:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT *
21:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT *
22:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	⁸ Query Match	Length	DB	ID	Description
1	91	49.7	309	12	AA131175	NS1_81-RLfcltag.
2	91	49.7	319	11	AA07945	NS181RLFAuth plasm
3	91	49.7	319	12	AA131176	NS1_81-RLFAuth. P
4	91	49.7	327	12	AA131177	NS1_81-RLFAuth + (
5	91	49.7	335	12	AA131178	NS1_81(NANP)4RLfau
6	91	49.7	335	12	AA131179	NS1_81(NVDP)4RLfau
7	91	49.7	411	9	AA683144	Sequence encoded b
8	91	49.7	412	7	AA660416	CS protein of mala
9	91	49.7	424	14	AA37796	RTS protein. Synth
10	91	49.7	424	14	AA37797	RTS* protein. Syn
11	89.5	48.9	33	22	AA63663	Peptide comprising

12	89.5	48.9	33	22	AAGC3516	A peptide which ma
13	89	48.6	19	22	AAW58951	Vaccine related MH
14	89	48.6	21	10	AAAP91504	Sequence of modifi
15	89	48.6	21	16	AAAR82586	Plasmodium falcipa
16	89	48.6	21	16	AAAR78920	Malaria circumspor
17	89	48.6	21	16	AAAR75955	P. falciparum CS p
18	89	48.6	21	16	AAAR70912	Malaria circumspor
19	89	48.6	21	17	AAW05612	Circumsporozoite h
20	89	48.6	21	18	AAW35440	T-cell stimulatory
21	89	48.6	21	20	AAW33252	Peptide derived fr
22	89	48.6	21	21	AAW80071	Pathogen derived T
23	89	48.6	21	21	AAW54553	T helper cell (Th)
24	89	48.6	21	21	AAW58777	Unidentified pepti
25	89	48.6	21	22	AAW9706	Plasmodium falcipa
26	89	48.6	21	22	AAW62478	Plasmodium falcipa
27	89	48.6	21	22	AAW64517	Plasmodium falcipa
28	89	48.6	21	22	AAW682859	Plasmodium falcipa
29	89	48.6	21	22	AAW89366	Plasmodium falcipa
30	89	48.6	21	22	AAW84447	Sequence of T help
31	89	48.6	21	22	AAW84457	Plasmodium falcipa
32	89	48.6	22	16	AAW82077	Malaria CS73 prote
33	89	48.6	33	15	AAW62715	LHRH-containing im
34	89	48.6	33	16	AAW83570	IGE CH4 region con
35	89	48.6	143	31	AAW495252	N6 polypeptide car
36	89	48.6	218	21	AAW19253	N10 polypeptide ca
37	89	48.6	240	21	AAW49254	N11 polypeptide ca
38	89	48.6	390	21	AAW49255	N19 polypeptide ca
39	87	47.5	412	9	AAW80835	Sequence encoded b
40	85	46.4	18	21	AAW49259	CD4+ T cell epitop
41	83	45.4	350	21	AAW70228	Recombinant vaccin
42	82	44.8	21	15	AAW65375	Helper T cell epit
43	81	44.3	21	21	AAW70283	Plasmodium falcipa
44	79	43.2	17	16	AAW78919	Malaria circumspor
45	79	43.2	17	16	AAW70911	Malaria circumspor

ALIGNMENTS

RESULT 1	
AA13175	
ID	AA13175 standard; Protein; 309 AA.
XX	
AC	AA13175;
XX	
DT	29-AUG-1991 (first entry)
XX	
DE	NS1_81-RUfdelta9.
XX	
KW	Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX	
XX	hybrid; influenza virus; non-structural protein 1; fusion.
OS	Plasmodium falciparum.
XX	
XX	Influenza virus (A/PR/8/34/).
XX	
Key	Location/Qualifiers
FT	1..81
FT	/label= "N-terminal of NS1
FT	/note= "Influenza virus nonstructural protein 1"
FT	82..87
FT	/label= synthetic linker
FT	88
FT	/label= artifact
FT	/note= "see comments"
FT	89..193
FT	/label= AAs 19-123 of CS protein
FT	/note= "Region 1 confg. flanking region less
FT	signal sequence"
FT	194..309
FT	/label= AAs 297-412 of CS protein
FT	/note= "Region II flanking region minus 9 N-term
FT	inal AAs"
XX	

Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGSPSLD-----EKRIAKMEKASSVFNNVNS 36
Db 43 IKPGSANKPKDLDYENDIEKKICKMEKCSSVFNNVNS 80

RESULT 13

ID Q9U0Q0 PRELIMINARY; PRT; 80 AA.
AC Q9U0Q0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D28;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GURP) from Brazil, Senegalese, and
Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ265948; CAB64173.1; -
DR InterPro; IPR003067; Circmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; Circmsprzoite.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1 80
FT NON_TER 1 80
SQ SEQUENCE 80 AA; 9046 MW; BA7689DI8F031C3E CRC64;

Query Match 49.7%; Score 91; DB 5; Length 80;
Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGSPSLD-----EKRIAKMEKASSVFNNVNS 36
Db 43 IKPGSANKPKDLDYENDIEKKICKMEKCSSVFNNVNS 80

RESULT 14

ID Q9U0P9 PRELIMINARY; PRT; 80 AA.
AC Q9U0P9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GURP) from Brazil, Senegalese, and
Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269951; CAB64176.1; -
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; Circmsprzoite.
DR SMART; SM00209; TSP1; 1.

Query Match 49.7%; Score 91; DB 5; Length 80;
Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9032 MW; ADEDE6F0E266AD98E CRC64;

Query Match 49.7%; Score 91; DB 5; Length 80;
Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGSPSLD-----EKRIAKMEKASSVFNNVNS 36
Db 43 IKPGSANKPKDLDYENDIEKKICKMEKCSSVFNNVNS 80

RESULT 15

ID Q9U0P8 PRELIMINARY; PRT; 80 AA.
AC Q9U0P8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M4;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
falciparum glutamate rich protein (GURP) from Brazil, Senegalese, and
Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ269957; CAB64237.1; -
DR InterPro; IPR003067; Circmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; Circmsprzoite.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1 80
FT NON_TER 1 80
SQ SEQUENCE 80 AA; 9073 MW; A8F404B8FB142B1E CRC64;

Query Match 49.7%; Score 91; DB 5; Length 80;
Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGSPSLD-----EKRIAKMEKASSVFNNVNS 36
Db 43 IKPGSANKPKDLDYENDIEKKICKMEKCSSVFNNVNS 80

Search completed: October 10, 2002, 16:10:04
Job time : 14 secs

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D4393, AND D4372:
RA de Stricker K., Vuust J., Jepsen S., Oeuvaray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269966; CAB64185.1; -.
DR EMBL: AJ269964; CAB64183.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9059 MW; A756D1FC41C1C21 CRC64;

Query Match 49.7%; Score 91; DB 5; Length 80;
Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKKIAKMEKASSVFNVVNS 36
Db 43 IKPGSANKPKDELNDYENDIEKKICKMEKCSSVFNVVNS 80
:||||: | ||||| ||||| ||||| |||||

RESULT 10
Q9TVN9 PRELIMINARY; PRT; 80 AA.
ID Q9TVN9;
AC Q9TVN9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MO, M6, M7, AND M1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvaray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269955; CAB64242.1; -.
DR EMBL: AJ269958; CAB64238.1; -.
DR EMBL: AJ269959; CAB64239.1; -.
DR EMBL: AJ269956; CAB64241.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9074 MW; A8F40C90DB1C033E CRC64;

Query Match 49.7%; Score 91; DB 5; Length 80;
Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKKIAKMEKASSVFNVVNS 36
Db 43 IKPGSANKPKDELNDYENDIEKKICKMEKCSSVFNVVNS 80
:||||: | ||||| ||||| ||||| |||||

RESULT 11
Q9TVN9 PRELIMINARY; PRT; 80 AA.
ID Q9TVN9;
AC Q9TVN9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BI;
RA de Stricker K., Vuust J., Jepsen S., Oeuvaray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269941; CAB64167.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9102 MW; A3283B70CEB50FDE CRC64;

Query Match 49.7%; Score 91; DB 5; Length 80;
```

```
AC Q9TVN9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B1893, B1804, B1853, B1870, B1872, B1881, AND B1882;
RA de Stricker K., Vuust J., Jepsen S., Oeuvaray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269977; CAB64196.1; -.
DR EMBL: AJ269971; CAB64190.1; -.
DR EMBL: AJ269972; CAB64191.1; -.
DR EMBL: AJ269973; CAB64192.1; -.
DR EMBL: AJ269974; CAB64193.1; -.
DR EMBL: AJ269975; CAB64194.1; -.
DR EMBL: AJ269976; CAB64195.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9074 MW; ADP10C8FC41C1924 CRC64;

Query Match 49.7%; Score 91; DB 5; Length 80;
Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKKIAKMEKASSVFNVVNS 36
Db 43 IKPGSANKPKDELNDYENDIEKKICKMEKCSSVFNVVNS 80
:||||: | ||||| ||||| ||||| |||||

RESULT 12
Q9U0Q4 PRELIMINARY; PRT; 80 AA.
ID Q9U0Q4;
AC Q9U0Q4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvaray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269941; CAB64167.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crcmsprzoite.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 1
FT NON_TER 80 80
SQ SEQUENCE 80 AA; 9102 MW; A3283B70CEB50FDE CRC64;

Query Match 49.7%; Score 91; DB 5; Length 80;
```

```

DR PRINTS: PRO1303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
SQ SEQUENCE 420 AA; 45155 MW; 3A85B92432C2893C CRC64;

Query Match
Best Local Similarity 57.9%; Score 97; DB 5; Length 420;
Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
DB 369 IKPGSANKPKDLDYANDIEKKICKMEKCSSVFNVNS 406

RESULT 6
Q9TW97 PRELIMINARY; PRT; 80 AA.
AC Q9TW97;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CIRCUMSPOROZITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA403, AND D4259;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269968; CAB64187.1; -.
DR EMBL: AJ269962; CAB64181.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PRO1303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 80
FT SEQUENCE 80 AA; 9000 MW; 036C7E319AFEL3B3 CRC64;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
DB 43 IKPGSANKPKDLDYANDIEKKICKMEKCSSVFNVNS 80

Query Match
Best Local Similarity 49.7%; Score 91; DB 5; Length 80;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

RESULT 7
Q9TW83 PRELIMINARY; PRT; 80 AA.
AC Q9TW83;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CIRCUMSPOROZITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA17, DA1, AND H1;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL: AJ269952; CAB64177.1; -.
DR EMBL: AJ269943; CAB64169.1; -.
DR EMBL: AJ269949; CAB64174.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PRO1303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 80
FT SEQUENCE 80 AA; 8989 MW; 0E769C8AC1030149 CRC64;

Query Match
Best Local Similarity 49.7%; Score 91; DB 5; Length 80;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
DB 43 IKPGSANKPKDLDYANDIEKKICKMEKCSSVFNVNS 80

RESULT 8
Q9TW76 PRELIMINARY; PRT; 80 AA.
AC Q9TW76;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CIRCUMSPOROZITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PALO ALTO, AND D50;
RA de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT "Sequence variation in the non-repeat region of the Plasmodium
RT falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT Burmese field isolates and from laboratory strains."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ269954; CAB64179.1; -.
DR EMBL: AJ269944; CAB64170.1; -.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PRO1303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1 80
FT SEQUENCE 80 AA; 9043 MW; 17282E319AE508B3 CRC64;

Query Match
Best Local Similarity 55.3%; Score 91; DB 5; Length 80;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
DB 43 IKPGSANKPKDLDYANDIEKKICKMEKCSSVFNVNS 80

RESULT 9
Q9TV00 PRELIMINARY; PRT; 80 AA.
AC Q9TV00;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CIRCUMSPOROZITE PROTEIN (FRAGMENT).
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;

```


Accession	Gene	Protein	Score	Length	Matches	Conservative	Mismatches	Indels	Gaps
AC_025729	01-NOV-1996 (TREMblrel. 01, Created)								
DT_01-NOV-1996 (TREMblrel. 01, Last sequence update)									
DT_01-DEC-2001 (TREMblrel. 19, Last annotation update)									
DE_CIRCUMSPOROZOITE PROTEIN.									
GN_CS.									
OS Plasmodium falciparum.									
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.									
XX NCBI_TaxID=5833;									
RM (1)									
RP SEQUENCE FROM N.A.									
RC STRAIN=SANTA LUCIA, SALI;									
RA Qari S.H., Lal A.A.;									
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.									
DR EMBL; U20959; AAA63153.1; -									
DR InterPro; IPR003067; Circmsporzoite.									
DR InterPro; IPR000884; TSP1.									
DR Pfam; PF00090; tsp_1; 1									
DR PRINTS; PRO1303; CIRCMSPRZOITE.									
DR SMART; SM00209; TSP1; 1.									
DR Malaria.									
SW SEQUENCE 408 AA; 43871 MW; A545BE517822515C CRC64;									
Query Match			55.7%;	Score 102;	DB 5;	Length 408;			
Best Local Similarity			60.5%;	Pred. No. 2e-06;					
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 1;									
QY 7 LRPGSSGPSLD-----EKKIAKMEKASSVENYVNS 36									
Db 357 IKPGSAGSKDELDYENDIEKKICKMEKCSVENYVNS 394									
RESULT 3									
ID 027325 PRELIMINARY; PRT: 436 AA.									
AC 027325.									
DT 01-NOV-1996 (TREMblrel. 01, Created)									
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)									
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)									
DE CIRCUMSPOROZOITE PROTEIN.									
GN_CSP.									
OS Plasmodium falciparum.									
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.									
OX NCBI_TaxID=5833;									
XX (1)									
RN SEQUENCE FROM N.A.									
RC STRAIN=837;									
RX MEDLINE=84250215; PubMed=6204383;									
RA Dane J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,									
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,									
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;									
RT "Structure of the gene encoding the immunodominant surface antigen on									
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";									
RT Science 225:593-599(1984).									
RL [2]									
RP SEQUENCE FROM N.A.									
RC STRAIN=837;									
RA Ia Cruz V.F.;									
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.									
RM [3]									
RN SEQUENCE FROM N.A.									
RC STRAIN=837;									
RA Jongwutinsin S., Tanabe K., Kanbara H.;									
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.									
DR EMBL; M83164; AAA29542.1; -									
DR EMBL; M83150; AAA29563.1; -									
DR EMBL; M83163; AAA29576.1; -									
DR InterPro; IPR003067; Circmsprzoite.									
DR InterPro; IPR000884; TSP1.									
DR Pfam; PF00090; tsp_1; 1									
DR PRINTS; PRO1303; CIRCMSPRZOITE.									

```

Query Match          55.7%; Score 102; DB 5; Length 436;
Best Local Similarity 60.5%; Pred. No. 2.1e-06;
Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

QY      7  LRPSSGSPSLD-----EKKIAMKEXASVFENVVN 36
Db      385 IKPGSAGSKSDLDYENDIEKKICKMEKCSSVFENVVN 422
          ::|||:| | | | | | | | | | | | | | | | | |
          ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
Q9U0Q2      PRELIMINARY; PRT; 79 AA.
AC      Q9U0Q2
Q9U0Q2:
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN      CS.
OS      Plasmodium falciparum.
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=5833;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=DD51;
RA      de Stricker K., Vuust J., Jepsen S., Oeuvray C., Theisen M.;
RT      "Sequence variation in the non-repeat region of the Plasmodium
RT      falciparum glutamate rich protein (GLURP) from Brazil, Senegalese, and
RT      Burmese field isolates and from laboratory strains.";
RL      Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ265945; CAB64243.1; -.
DR      InterPro: IPR003067; Crcmsprzoite.
DR      InterPro: IPR000884; TSP1.
DR      Pfam: PF00090; tsp_1; 1.
DR      PRINTS; PRO1303; CRCMSPRZOITE.
DR      SMART; SM00209; TSP1, 1.
DR      NON_TER 1
FT      NON_TER 79
FT      TER 79
SQ      SEQUENCE 79 AA; 8893 MW; 56BFA76D859BA16 CRC64;

Query Match          53.6%; Score 98; DB 5; Length 79;
Best Local Similarity 59.5%; Pred. No. 1.3e-06;
Matches 22; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

QY      7  LRPSSGSPSLD-----EKKIAMKEXASVFENVVN 35
Db      43  IKPGSAGSKSDLDYENDIEKKICKMEKCSSVFENVVN 79
          ::|||:| | | | | | | | | | | | | | | | | |
          ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
Q25838      PRELIMINARY; PRT; 420 AA.
AC      Q25838
Q25838:
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      CIRCUMSPOROZOITE PROTEIN.
GN      CSP.
OS      Plasmodium falciparum.
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=5833;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=835B;
RX      MEDLINE=95077069; PubMed=7985759;
RA      Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT      "Allelic variation in the circumsporozoite protein of Plasmodium
RT      falciparum from Thai field isolates.";
RL      Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR      EMBL; M83161; AAA29574.1; -.
DR      InterPro: IPR003067; Crcmsprzoite.
DR      InterPro: IPR000884; TSP1.
DR      Pfam: PF00090; tsp_1; 1.

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 14 Seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-16
Perfect score: 183
Sequence: 1 XHMSYGLRPGSSGSLDEKKIAKMEKASVFNVNS 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SPREMBL_19:*
 - 2: sp_archaea:*
 - 3: sp_bacteria:*
 - 4: sp_fungi:*
 - 5: sp_human:*
 - 6: sp_invertebrate:*
 - 7: sp_mammal:*
 - 8: sp_mhc:*
 - 9: sp_organelle:*
 - 10: sp_phage:*
 - 11: sp_plant:*
 - 12: sp_prodent:*
 - 13: sp_virus:*
 - 14: sp_vertebrate:*
 - 15: sp_unclassified:*
 - 16: sp_virus:*
 - 17: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	102	55.7	117	5	Q25797	Q25797 plasmodium
2	102	55.7	408	5	Q25729	Q25729 plasmodium
3	102	55.7	436	5	Q27325	Q27325 plasmodium
4	98	53.6	79	5	Q9U002	Q9U002 plasmodium
5	97	53.0	420	5	Q25838	Q25838 plasmodium
6	91	49.7	80	5	Q9TW97	Q9TW97 plasmodium
7	91	49.7	80	5	Q9TW83	Q9TW83 plasmodium
8	91	49.7	80	5	Q9TW76	Q9TW76 plasmodium
9	91	49.7	80	5	Q9TW60	Q9TW60 plasmodium
10	91	49.7	80	5	Q9TW99	Q9TW99 plasmodium
11	91	49.7	80	5	Q9TW99	Q9TW99 plasmodium
12	91	49.7	80	5	Q9U004	Q9U004 plasmodium
13	91	49.7	80	5	Q9U000	Q9U000 plasmodium
14	91	49.7	80	5	Q9U0P9	Q9U0P9 plasmodium
15	91	49.7	80	5	Q9U0P8	Q9U0P8 plasmodium
16	91	49.7	80	5	Q9U0P7	Q9U0P7 plasmodium

17	91	49.7	80	5	Q9U0P6	Q9U0P6 plasmodium
18	91	49.7	80	5	Q9U0P5	Q9U0P5 plasmodium
19	91	49.7	80	5	Q9U0P4	Q9U0P4 plasmodium
20	91	49.7	80	5	Q9U0P3	Q9U0P3 plasmodium
21	91	49.7	80	5	Q9U0P2	Q9U0P2 plasmodium
22	91	49.7	80	5	Q9U0P1	Q9U0P1 plasmodium
23	91	49.7	115	5	Q9TW77	Q9TW77 plasmodium
24	91	49.7	115	5	Q25835	Q25835 plasmodium
25	91	49.7	115	5	Q25836	Q25836 plasmodium
26	91	49.7	115	5	Q25839	Q25839 plasmodium
27	91	49.7	117	5	Q25794	Q25794 plasmodium
28	91	49.7	117	5	Q25795	Q25795 plasmodium
29	91	49.7	117	5	Q25796	Q25796 plasmodium
30	91	49.7	393	5	Q99255	Q99255 plasmodium
31	91	49.7	424	5	Q27425	Q27425 plasmodium
32	91	49.7	424	5	Q99256	Q99256 plasmodium
33	91	49.7	432	5	Q27246	Q27246 plasmodium
34	91	49.7	432	5	Q25827	Q25827 plasmodium
35	90	49.2	80	5	Q9TW01	Q9TW01 plasmodium
36	90	49.2	80	5	Q9TW00	Q9TW00 plasmodium
37	90	49.2	115	5	Q9U934	Q9U934 plasmodium
38	90	49.2	115	5	Q25837	Q25837 plasmodium
39	90	49.2	416	5	Q25829	Q25829 plasmodium
40	90	49.2	420	5	Q25831	Q25831 plasmodium
41	90	49.2	436	5	Q25828	Q25828 plasmodium
42	90	49.2	442	5	Q25830	Q25830 plasmodium
43	90	49.2	452	5	Q25834	Q25834 plasmodium
44	88	48.1	80	5	Q9U0Q3	Q9U0Q3 plasmodium
45	87	47.5	80	5	Q9U0Q1	Q9U0Q1 plasmodium

ALIGNMENTS

RESULT 1

ID Q25797 PRELIMINARY: PRT: 117 AA.

AC Q25797

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DR 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5633;

RN [1]

RP SEQUENCE FROM N.A.

RA Doolan D.L., Saul A., Good M.F.;

RT "Geographically restricted heterogeneity of the plasmodium falciparum circumsporozone protein: relevance for vaccine development.";

RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

DR EMBL; M7205; AAA29519.2; -

DR InterPro: IPR003067; Circumsporozone.

DR InterPro: IPR000884; TSP1.

DR Pfam: PF000090; TSP_1; 1.

DR PRINTS: PRO1303; CIRCUMSPOROITE.

DR SMART; SM00209; TSP1; 1.

FT NON_TER 1 117

FT NON_TER 117 117

SQ - SEQUENCE 117 AA: 13043 MW: ODA711D86C0B03C1 CMC64;

Query Match 55.7%; Score 102; DB 5; Length 117;

Best Local Similarity 60.5%; Pred. No. 5.1e-07;

Matches 23; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

QY 7 LRPSSGSLD-----EKIKMEKASVFNVNS 36

Db 66 IKPGSAGSKDELVDYNDIEKKIKMEKCSVFNVNS 103

RESULT 2

ID Q25729 PRELIMINARY: PRT: 408 AA.

```

OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Scandentia; Tupaiidae; Tupata.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus; PubMed=8921350;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGnRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GnRH gene expression in a
RT placental mammal."
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U63326; AAB16837.1; -
DR InterPro: IPR002012; GnRH.
DR InterPro: IPR004079; Gonadolibertin.
DR Pfam: PF00446; GnRH; 1.
DR PRINTS: PR01541; GONADOLIBERTIN.
DR PROSITE: PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW placenta; signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 1 24 PROGONADOLIBERTIN I.
FT PEPTIDE 24 33 GONADOLIBERTIN I.
FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 24 24 PYROGLUTAMINE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 92 AA; 10197 MW; 4FDBFC58CF5F63B CRC64;
Query Match 31.7%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33

```

```

RESULT 15
CSP_PLABE
ID CSP_PLABE STANDARD; PRT; 339 AA.
AC P06915;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87089740; PubMed=2432395;
RA Eichinger D.J., Arnot D.E., Tam J.P., Nussenzweig V., Enea V.;
RT "Circumsporozoite protein of Plasmodium berghei: gene cloning and

```

```

RT identification of the immunodominant epitopes."
RL Mol. Cell. Biol. 6:3965-3972(1986).
CC -1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
CC SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
CC VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14135; AAA29577.1; -
DR PIR: A25083; 02ZOMB.
DR InterPro: IPR003067; Circmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF000090; tsp.1; 1.
DR PRINTS: PR01303; CIRCMSPROZITE.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; signal.
FT SIGNAL 1 23 PROBABLE.
FT CHAIN 24 339 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 93 196 13 X 8 AA TANDDEM REPEATS.
FT DOMAIN 206 238 16 X 2 AA TANDDEM REPEATS OF P-Q.
SQ SEQUENCE 339 AA; 37138 MW; E8068A6D1D9551B CRC64;
Query Match 31.7%; Score 58; DB 1; Length 339;
Best Local Similarity 37.1%; Pred. No. 0.69;
Matches 13; Conservative 9; Mismatches 7; Indels 6; Gaps 1;
OY 8 RFGSSGSPSLD-----EKKIAKMEKASSVFNVVNS 36
DB 291 RKGSKKKAEDLTLEDIDTEICKMDKCSIFNIVSN 325

```

Search completed: October 10, 2002, 16:06:35
Job time : 5.69231 secs

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X01059; CAA3526.1; -
DR EMBL: M12578; AAA35916.1; -
DR EMBL: X15215; CAA33285.1; -
DR PIR: A01410; RHUG.
DR PIR: A26173; A26173.
DR PIR: S05308; S05308.
DR MIM: 152760; -
DR InterPro: IPR002012; GNRH.
DR InterPro: IPR004079; GonadolibertinI.
DR Pfam: PF00446; GNRH; 1.
DR PRINTS: PRO1541; GONADOLIBERTIN.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT CONFLICT 16 16
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 31.7%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
    |||||
Db 25 HWSYGLRPG 33

RESULT 13
GONL_RAT STANDARD; PRT; 92 AA.
ID GONL_RAT
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Progonaolibertin I precursor [Contains: Gonadolibertin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Lulberin I); Prolactin release-inhibiting factor
DE 1].
GN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat."
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone. SH locus: structure and
RT hypothalamic expression."
RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;

```

```

RX MEDLINE=93105480; PubMed=1468115;
RA Water C.C., Marchetti B., Leboeuf R.D., Blacklock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA."
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus."
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S50870; AAB24572.1; -
DR EMBL: M12579; AAA41263.1; -
DR EMBL: M31670; AAA41264.1; -
DR EMBL: M15527; AAA42141.1; ALT-SEQ.
DR EMBL: M15529; AAA42139.1; -
DR EMBL: M15528; -; NOT_ANNOTATED_CDS.
DR PIR: B26173; RHRTG.
DR PIR: A48410; A48410.
DR InterPro: IPR002012; GNRH.
DR InterPro: IPR004079; GonadolibertinI.
DR Pfam: PF00446; GNRH; 1.
DR PRINTS: PRO1541; GONADOLIBERTIN.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match 31.7%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
    |||||
Db 25 HWSYGLRPG 33

RESULT 14
GONL_TUPGB STANDARD; PRT; 92 AA.
ID GONL_TUPGB
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonaolibertin I precursor [Contains: Gonadolibertin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Lulberin I); GNRH-associated peptide I].
GN GNRH1 OR GNRH.

```

OY 2 HWSYGLRPG 10
 DB 23 HWSYGLRPG 31

RESULT 11

ID	GNL_PIG	STANDARD:	PRT:	91 AA.
AC	P49921;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)]			
DE	(Luteinizing hormone releasing hormone I) (Gonadotropin releasing hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].			
GN	GNRH OR GNRH.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Hypothalamus;			
RA	Weesner G.D., Mattern R.L., Becker B.A.;			
RL	Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE OF 24-33.			
RA	MEDLINE=72114303; PubMed=4946067;			
RA	Baba Y., Matsuo H., Schally A.V.;			
RT	"Structure of the porcine LH- and FSH-releasing hormone. II.			
RT	Confirmation of the proposed structure by conventional sequential analyses."			
RL	Biochem. Biophys. Res. Commun. 44:459-463(1971).			
RN	[3]			
RP	SYNTHESIS OF GONADOLIBERIN.			
RA	MEDLINE=72065376; PubMed=4942726;			
RA	Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;			
RT	"Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase method."			
RL	Biochem. Biophys. Res. Commun. 45:822-827(1971).			
RN	[4]			
RP	SYNTHESIS OF GONADOLIBERIN.			
RA	MEDLINE=72117544; PubMed=4946275;			
RA	Baba Y., Arimura A., Schally A.V.;			
RT	"On the tryptophan residue in porcine LH and FSH-releasing hormone."			
RL	Biochem. Biophys. Res. Commun. 45:483-487(1971).			
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L32864; AAA31066.1; -.			
DR	PIR; A01411; RHPGG.			
DR	InterPro; IPR002012; GNRH.			
DR	InterPro; IPR004079; GonadoliberinI.			
DR	Pfam; PF00446; GNRH; 1.			
DR	PRINTS; PR01541; GONADOLIBRNT.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;			
FT	CHAIN 1 23			
FT	SIGNAL 24 91			
FT	PEPTIDE 24 33			
		PROGONADOLIBERIN I.		
		GONADOLIBERIN I.		

FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT MOD_RES 24 24 ACTIVITY.
 FT MOD_RES 33 33 PYROLIDONE CARBOXYLIC ACID.
 SO SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64; AMIDATION (G-34 PROVIDE AMIDE GROUP).

Query Match
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 DB 25 HWSYGLRPG 33

RESULT 12

ID	GNL_HUMAN	STANDARD:	PRT:	92 AA.
AC	P01148;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)]			
DE	(Luteinizing hormone releasing hormone I) (Gonadotropin releasing hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated peptide I].			
GN	GNRH OR GNRH OR LHRH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=8936682; PubMed=2671939;			
RA	Hayflick J.S., Adelman J.P., Seeburg P.H.;			
RT	"The complete nucleotide sequence of the human gonadotropin-releasing hormone gene."			
RL	Nucleic Acids Res. 17:6403-6403(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=86094338; PubMed=2867548;			
RA	Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;			
RT	"Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat."			
RL	Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=85012739; PubMed=6090951;			
RA	Seeburg P.H., Adelman J.P.;			
RT	"Characterization of cDNA for precursor of human luteinizing hormone releasing hormone."			
RL	Nature 311:666-668(1984).			
RN	[4]			
RP	SEQUENCE OF 24-33.			
RA	MEDLINE=83126573; PubMed=6760865;			
RA	Tan L., Rousseau P.;			
RT	"The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in the human placenta."			
RL	Biochem. Biophys. Res. Commun. 109:1061-1071(1982).			
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs), Lutrepulse or Lutrelaf (Perring Pharmaceuticals) and Relisorm (Sergo).			
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its			

```

FT  NON_T2R      1      1
FT  SIGNAL        <1      5
FT  CHAIN         6      >67
FT  PEPTIDE       6      15
FT  ACT_SITE     19      >67
FT  MOD_RES       8      8
FT  MOD_RES       6      6
FT  MOD_RES      15      15
FT  NON_TER      67      67
SQ  SEQUENCE     67 AA: 7573 MW: 505394DMA261A3F2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 58; DB 1; Length 67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 HWSYGLRPG 10
Db  7 HWSYGLRPG 15

RESULT 9
GONI_XENLA
ID  GONI_XENLA  STANDARD;  PRT;  89 AA.
AC  P45656;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
DE  (LH-RH) (Luliberin I)
OS  Xenopus laevis (African clawed frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC  Xenopodidae; Xenopus.
OX  NCBI_TaxID=8355;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Forebrain;
RX  MEDLINE=94185563; PubMed=8137750;
RA  Hayes W.P., Wray S., Batley J.F.;
RT  "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT  mammalian-like expression pattern and conserved domains in
RT  GNRH-associated peptide, but brain onset is delayed until
RT  metamorphosis."
RL  Endocrinology 134:1835-1844(1994).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: L28040; AAA49728.1; -
DR  InterPro: IPR002012; GNRH.
DR  InterPro: IPR004079; GonadoliberinI.
DR  Pfam: PF00446; GNRH; 1.
DR  PRINTS: PR01541; GONADOLIBERIN.
DR  PROSITE: PS00473; GNRH; 1.
KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW  Signal.
FT  SIGNAL        1      23
FT  CHAIN         24      89
FT  PEPTIDE       24      33
FT  ACT_SITE     37      89
FT  MOD_RES      37      89
FT  SEQUENCE     37      85
Query Match
Best Local Similarity 100.0%; Score 58; DB 1; Length 67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

FT  MOD_RES      24      24
FT  MOD_RES      33      33
SQ  SEQUENCE     89 AA: 10246 MW: 6FAF36FBAE0D4284 CRC64;

Query Match
Best Local Similarity 100.0%; Score 58; DB 1; Length 89;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 HWSYGLRPG 10
Db  25 HWSYGLRPG 33

RESULT 10
GONI_MOUSE
ID  GONI_MOUSE  STANDARD;  PRT;  90 AA.
AC  P13562;
DT  01-JAN-1990 (Rel. 13, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE  (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE  hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor
DE  I].
GN  GNRH1 OR GNRH.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=8706928; PubMed=3024317;
RA  Mason A.T., Haylick J.S., Zoeller R.T., Young W.S. III,
RA  Phillips H.S., Nikolic K., Seeburg P.H.;
RT  "A deletion truncating the gonadotropin-releasing hormone gene is
RT  responsible for hypogonadism in the hpg mouse."
RL  Science 234:1366-1371(1986).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. IT STIMULATES
CC  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC  HORMONES.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: M14872; AAA37717.1; -
DR  MGD; MGI:95789; GnRh.
DR  InterPro: IPR002012; GNRH.
DR  InterPro: IPR004079; GonadoliberinI.
DR  Pfam: PF00446; GNRH; 1.
DR  PRINTS: PR01541; GONADOLIBERIN.
DR  PROSITE: PS00473; GNRH; 1.
KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW  Placenta; Signal.
FT  SIGNAL        1      21
FT  CHAIN         22      90
FT  PEPTIDE       22      31
FT  ACT_SITE     35      90
FT  MOD_RES      24      24
FT  MOD_RES      22      22
FT  MOD_RES      31      31
SQ  SEQUENCE     90 AA: 10337 MW: 1C0766FA826E4D9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 58; DB 1; Length 90;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR DR EMBL; U02517; AAA03433.1; -.
DR PIR; A93780; RHSHG.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT FT NON_TER 1 1
FT CHAIN 1 >61 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >61 GNRR-ASSOCIATED PEPTIDE I.
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-I1 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
SO SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;
OY Query Match 33.6%; Score 61.5; DB 1; Length 61;
Db Best Local Similarity 42.9%; Pred. No. 0.029; Mismatches 2; Indels 15; Gaps 2;
Matches 15; Conservative 3;
OY 2 HWSYGLRPGSSPSLDEKKIAKMEASVFNVNS 36
Db 2 HWSYGLRPGG-----RRMAK-----NVIDS 21
| | | | | | | | | | : | | | | |
RESULT 7
GONI_MESAU STANDARD: PRT; 63 AA.
AC ID GONI_MESAU STANDARD: PRT; 63 AA.
AC 009163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Gonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I); GNRR-associated peptide I]
DE (Fragment).
DE GN GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxId=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR DR EMBL; U91938; AAB51302.1; -.
DR InterPro; IPR002012; GnRH.

```

KW	Pfam; PF00446; GNRH; 1.
DR	PROSITE; PS00473; GNRH; 1.
KM	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
FT	Placenta.
FT	NON_TER
FT	CHAIN
FT	PEPTIDE
FT	PEPTIDE
FT	ACT_SITE
FT	MOD_RES
FT	MOD_RES
FT	MOD_RES
FT	NON_TER
SO	SEQUENCE
Oy	Query Match
Db	Best Local Similarity
	Matches
	2 HWSYGLRPGSS-----GPSLDE 18
	2 HMSYGLRPGGRNMRERLDSPQE 24
	RESULT 8
ID	GON1_MACMU
AC	GON1_MACMU STANDARD; PRT; 67 AA.
DT	P55247:
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Progonadoliblerin I precursor [Cntrlas: Gonadoliblerin I (LHRH I) (humanizing hormone releasing hormone I) (Gonadotropin releasing hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I] (Fragment).
DE	GNRH1 OR GNRH OR LHRH.
GN	Maca mulatta (rhesus macaque).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Euteria; Primates; Carnivora; Cercopithecoidea;
OC	Cercopithecinae; Macaca.
OX	NCBI_Taxid=9544;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Hypothalamus;
RA	MEDLINE=95124501; PubMed=7545971;
RA	Ma Y.Y., Costa M.E., Ojeda S.R.;
RT	"Developmental expression of the genes encoding transforming growth factor alpha and its receptor in the hypothalamus of female rhesus macaques."
RL	Neuroendocrinology 60:346-359(1994).
CC	- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.
CC	- SUBCELLULAR LOCATION: Secreted.
CC	- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; S75918; AAB33096.1; -
DR	InterPro; IPRO02012; GNRH.
DR	Pfam; PF00446; GNRH; 1.
DR	PROSITE; PS00473; GNRH; 1.
KM	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Signal.

```

DT 01-JUN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate 14 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from Thailand."
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19752; AAA2955.1; -.
DR PIR; A54533; A54533.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 424 CIRCUMSPOROZITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 424 AA; 45610 MW; 710AB1A238786C93 CRC64;

Query Match 49.7%; Score 91; DB 1; Length 424;
Best Local Similarity 55.3%; Pred. NO. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
Db 373 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNS 410

RESULT 5
CSP_PLAFW STANDARD; PRT; 442 AA.
AC P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J., Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -1- FUNCTION: THE CIRCUMSPOROZITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZITE (THE INFECTIVE STAGE OF THE

```

```

CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15505; AAA2955.1; -.
DR PIR; A54529; A54529.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 CIRCUMSPOROZITE PROTEIN.
FT CHAIN 17 442
FT DOMAIN 130 320 47 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

Query Match 49.2%; Score 90; DB 1; Length 442;
Best Local Similarity 55.3%; Pred. NO. 1.9e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNS 36
Db 391 IKPGSADKPKQDLVDYENDIEKKICKMEKCSSVFNVNS 428

RESULT 6
GN1_SHEEP STANDARD; PRT; 61 AA.
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LHRH I) (luteinizing hormone releasing hormone I) (Gonadotropin releasing hormone I) (GNRH I) (luliberin I); GNRH-associated peptide I] (Fragment).
DE (Fragment).
DE DE
CN GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN=WESTERN RANGE; TISSUE=Hypothalamus;
RA Rodriguez R.E., Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-10.
RX MEDLINE=72094314; PubMed=4550508;
RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J., Fellows R., Blackwell R., Vale W., Guillemin R.;
RT "Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor (LRF) (LH-hypothalamus-irf-gas chromatography-mass spectrometry-decapeptide-Bdman degradation)."
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -1- THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.

```


RESULT 2

CSP_PLAFO STANDARD; PRT; 397 AA.

ID CSP_PLAFO
AC P19597; Q25798; 17, Created)
DT 01-FEB-1991 (Rel. 17, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5843;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=89345189; PubMed=2668895;
RA Campbell J.R.;
RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria
candidate vaccine antigen.";
RL Nucleic Acids Res. 17:5854-5854(1989).
[2]
REVIEWS.
RP Campbell J.R.;
RN Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RX MEDLINE=92155298; PubMed=1346766;
RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
Thomas A.W., Bagar S., Cochran M.A., Thanassi J., Levine M.M.,
Hackett C.S.;
RT "Plasmodium falciparum: in vitro characterization and human
infectivity of a cloned line.";
RL Exp. Parasitol. 74:159-168(1992).
[4]
SEQUENCE FROM N.A.
RX MEDLINE=89364998; PubMed=2671723;
RA Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;
RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
isolate used in malaria vaccine trials.";
RL Mol. Biochem. Parasitol. 35:185-190(1989).
-1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
SURFACE ANTIGEN ON THE SPOROZOITE (THE INJECTIVE STAGE OF THE
MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: X15363; CAA33421.1; -
DR EMBL: M83886; AAA29521.1; -
DR EMBL: M22982; AAA29527.1; -
DR PIR: S05428; S05428.
DR PIR: A45527; A45527.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.
FT CONFLICT 194 194 A -> ANPANPNA (TN REF. 4).
SQ SEQUENCE 397 AA; 42646 MM; 9E81146F59EBCA3 CRC64;

Query Match 49.7%; Score 91; DB 1; Length 397;
Best Local Similarity 55.3%; Pred. No. 1.2e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPGSSGPSLD-----EKKIAKMERASSVFNVVNS 36

Db 346 IKPGSANKPKDELVDYANDIEKKICKMEKSCSVFNVVNS 383

RESULT 3

CSP_PLAFA STANDARD; PRT; 412 AA.

ID CSP_PLAFA
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5833;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
Sanders G.S., Reddy E.P., Digs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
-1- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
SURFACE ANTIGEN ON THE SPOROZOITE (THE INJECTIVE STAGE OF THE
MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
VERTEBRATE HOST).
CC -1- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -1- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: K02194; AAA29524.1; -
DR PIR: A03388; OZG0AF.
DR InterPro: IPR003067; Crcmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
SQ SEQUENCE 412 AA; 44420 MM; 1EEED3DE90965F8 CRC64;

Query Match 49.7%; Score 91; DB 1; Length 412;
Best Local Similarity 55.3%; Pred. No. 1.3e-05;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPGSSGPSLD-----EKKIAKMERASSVFNVVNS 36

Db 361 IKPGSANKPKDELVDYANDIEKKICKMEKSCSVFNVVNS 398

RESULT 4

CSP_PLAFT STANDARD; PRT; 424 AA.

ID CSP_PLAFT
AC P13814;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 4.69231 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: us-09-848-834a-16
Sequence: 1 XHMSYGLRPGSSGSPSLDEKRIAKMEKASVFNVNS 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	53.0	388	1	CSP_PLARE
2	91	49.7	397	1	CSP_PLAFO
3	91	49.7	412	1	CSP_PLAFA
4	91	49.7	424	1	CSP_PLAFT
5	90	49.2	442	1	CSP_PLAFW
6	61.5	33.6	61	1	GONI_SHEEP
7	59	32.2	63	1	GONI_MESAU
8	58	31.7	89	1	GONI_MACMU
9	58	31.7	90	1	GONI_XENIA
10	58	31.7	90	1	GONI_MOUSE
11	58	31.7	91	1	GONI_PIG
12	58	31.7	92	1	GONI_HUMAN
13	58	31.7	92	1	GONI_RAT
14	58	31.7	92	1	GONI_TUPGB
15	58	31.7	339	1	CSP_PLARE
16	58	31.7	347	1	CSP_PLARE
17	56.5	30.9	90	1	GONI_DICLA
18	55.5	30.3	89	1	GONI_PORNO
19	55	30.1	94	1	GONI_HAPBU
20	55	30.1	367	1	CSP_PLAFO
21	54	29.5	10	1	GONI_ALAMI
22	54	29.5	92	1	GONI_CHICK
23	53.5	29.2	90	1	GONI_HAPBU
24	52.5	28.4	487	1	MMSA_BACSU
25	52	28.4	95	1	GONI_MORSA
26	52	28.4	95	1	GONI_PAGMA
27	52	28.4	95	1	GONI_SPAUV
28	52	28.4	99	1	GONI_DICLA
29	51.5	28.1	80	1	GONI_CLAGA
30	51.5	28.1	444	1	TIG_MYCE
31	51	27.9	393	1	CSP_PLAHR
32	51	27.9	429	1	CSP_PLAMA
33	51	27.9	959	1	MSH1_YEAST

34	50.5	27.6	90	1	GONI_PAGMA	P51921	pagrus majo
35	50.5	27.6	90	1	GONI_SPAUV	P51923	sparus aura
36	50.5	27.6	222	1	GTA2_CHICK	Q08393	gallus galli
37	50	27.3	90	1	GONI_RANDY	Q91au2	rana dybows
38	50	27.3	223	1	GTA1_RABIT	Q08863	oryctolagus
39	50	27.3	590	1	GTA1_BACSU	P45740	baillus su
40	50	27.3	595	1	THIC_BACHD	Q9kbj4	baillus ha
41	49.5	27.0	482	1	GABD_ECOLI	P25326	eschertichia
42	49	26.8	92	1	GONI_CAVPO	O54713	cavia porce
43	49	26.8	246	1	YSW4_CAEEL	Q10019	caenorhabdi
44	49	26.8	342	1	TRPD_BUCAI	P57367	buchnera ap
45	49	26.8	821	1	LIMS_CAEEL	P45970	caenorhabdi

ALIGNMENTS

RESULT 1

CSP_PLARE ID CSP_PLARE STANDARD; PRT; 388 AA.

AC P26694; DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Circumsporozoite protein precursor (CS).

OS Plasmodium reichenowi.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5854;

RN [1]

RP MEDLINE-91201303; PubMed=2016283;

RA Lal A.A., Goldman I.F.;

RT "Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria parasite evolutionarily related to the human

RT malaria parasite Plasmodium falciparum.";

RL J. Biol. Chem. 266:6686-6689 (1991).

CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE

CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE

CC VERTERATE HOST).

CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR

CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES

CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.

CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M60972; AAA29561.1; -.

DR PIR: A39756; A39756.

DR InterPro: IPR003067; Circmsprzoite.

DR InterPro: IPR000884; TSP1.

DR Pfam: PF00090; tsp_1; 1.

DR PRINTS: PR01303; CIRCMSPRZOITE.

DR SMART: SM00209; TSP1; 1.

KM Malaria: Sporozoite; Repeat: Signal.

FT SIGNAL 1 16 PROBABLE.

FT CHAIN 17 38 CIRCUMSPOROZOITE PROTEIN.

FT DOMAIN 120 267 62 x 4 AA TANDEM REPEATS OF N-A-N-P.

SQ SEQUENCE 388 AA; 42245 MW; C031EFBE2E35604 CRC64.

Query Match Score 97; DB 1; Length 388;
Best Local Similarity 57.9%; Pred. No. 1.6e-06;
Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRRGSSPSLD-----EKRIAKMEKASVFNVNS 36
Db 337 IKPGSAGKPKDLDYENDLEKRIKMEKASVFNVNS 374

R:Seeburg, P.H.; Adelman, J.P.
 Nature 311, 666-668, 1984
 A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing h
 A:Reference number: A93342; M0ID:85012739
 A:Accession: A93342
 A:Molecule type: mRNA
 A:Residues: 1-15, 'S', 17-92 <SEE>
 A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
 A:Experimental source: placenta
 R:Tan, L.; Rousseau, P.
 Blochem. Biophys. Res. Commun. 109, 1061-1071, 1982
 A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
 A:Reference number: A90108; M0ID:83126573
 A:Accession: A90108
 A:Molecule type: protein
 A:Residues: 24-33 <TAN>
 A:Experimental source: placental trophoblasts
 R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterda
 FEBS Lett. 346, 203-206, 1994
 A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
 A:Reference number: S45718; M0ID:94283597
 A:Contents: annotation; degradation pathway of synthetic hormone
 C:Genetics:
 A:Gene: GDB:GNRH; LHRH; GRH
 A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
 A:Map position: 8p21-8p11.2
 A:Introns: 47/3; 79/73
 C:Function:
 A:Description: gonadolibetin stimulates pituitary secretion of lutropin and follitropin
 A:Note: gonadolibetin-associated protein may have prolactin release inhibiting activity
 C:Superfamily: gonadolibetin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-92/Product: progonaolibetin #status predicted <PGN>
 F:24-33/Product: gonadolibetin #status experimental <MAT>
 F:37-92/Product: gonadolibetin-associated protein #status predicted <GAP>
 F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 31.7%; Score 58; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.86;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 Db 25 HWSYGLRPG 33
 |||||

RESULT 12
 gonadolibetin precursor - rat
 N:Alternate names: gonadolibetin-associated protein (GAP); gonadotropin releasing hormo
 N:Contains: gonadolibetin; prolactin release-inhibiting factor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-1988 #sequence; revision 31-Mar-1988 #text-change 18-Jun-1999
 C:Accession: A40147; B26173; A48410
 R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
 Mol. Endocrinol. 3, 1257-1262, 1989
 A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
 A:Reference number: A40147; M0ID:89384661
 A:Accession: A40147
 A:Molecule type: DNA
 A:Residues: 1-92 <BDN>
 A:Cross-references: GB:M31670; NID:g20447; PIDN:AAA1264.1; PID:g20448
 R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonado
 A:Reference number: A94090; M0ID:86094338
 A:Accession: B26173
 A:Molecule type: mRNA
 A:Residues: 1-92 <ADE>
 A:Cross-references: GB:M12579; NID:g20445; PIDN:AAA1263.1; PID:g20446
 R:Maier, C.C.; Marchetti, B.; Leboeuf, R.D.; Bialock, J.E.

Cell. Mol. Neurobiol. 12, 447-454, 1992
 A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing horm
 A:Reference number: A48410; M0ID:93105480
 A:Accession: A48410
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-92 <MAI>
 A:Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
 A:Experimental source: thymus
 A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:121083)
 C:Genetics:
 A:Introns: 47/3; 79/73
 C:Function:
 A:Description: stimulates pituitary secretion of lutropin and follitropin
 A:Note: gonadolibetin-associated protein may have prolactin release inhibiting activi
 C:Superfamily: gonadolibetin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-92/Product: progonaolibetin #status predicted <PGN>
 F:24-33/Product: gonadolibetin #status predicted <GLN>
 F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
 F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
 F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 31.7%; Score 58; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.86;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 Db 25 HWSYGLRPG 33
 |||||

RESULT 13
 OZOMB
 circumsporozoite protein precursor - Plasmodium berghei (strain NK65)
 N:Alternate names: sporozoite surface antigen
 C:Species: Plasmodium berghei
 C:Date: 30-Sep-1987 #sequence; revision 28-Jul-1995 #text-change 16-Jul-1999
 C:Accession: A44948; A25083; SL3446
 R:Lanar, D.E.
 Mol. Biochem. Parasitol. 39, 151-154, 1990
 A:Title: Sequence of the circumsporozoite gene of Plasmodium berghei ANKA clone and N
 A:Reference number: A44948; M0ID:90158693
 A:Accession: A44948
 A:Molecule type: DNA
 A:Residues: 1-332 <LAN>
 A:Cross-references: GB:M28887
 R:Eichinger, D.J.; Arnol, D.E.; Tam, J.P.; Nussenzweig, V.; Enea, V.
 Mol. Cell. Biol. 6, 3965-3972, 1986
 A:Title: Circumsporozoite protein of Plasmodium berghei: gene cloning and identificat
 A:Reference number: A25083; M0ID:87089740
 A:Accession: A25083
 A:Molecule type: DNA
 A:Residues: 1-26, 'T', 28-68, 'PMIRR', 75-126, 'P', 128-134, 'PPNANDP', 135-332 <EIC>
 A:Cross-references: GB:M14135; NID:g160245; PIDN:AAA29577.1; PID:g160246
 R:Weber, J.L.; Egan, J.E.; Lyon, J.A.; Wirtz, R.A.; Charoenyit, Y.; Maloy, W.L.; Hock
 Exp. Parasitol. 63, 295-300, 1987
 A:Title: Plasmodium berghei: cloning of the circumsporozoite protein gene.
 A:Reference number: SL3446; M0ID:87218962
 A:Accession: SL3446
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 61-122, 'A', 124-332 <WEB>
 A:Cross-references: GB:M25445; NID:g160177; PIDN:AAA29531.1; PID:g160178
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, t
 obic membrane-anchoring sequence.
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-332/Product: circumsporozoite protein #status predicted <MAT>
 F:94-189/Region: 8-residue repeats
 F:199-230/Region: 2-residue repeats

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
 C:Accession: A93780; A01411
 R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; Bl
 Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
 A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing facto
 A:Reference number: A93780; MUID:72094314
 A:Accession: A93780
 A:Molecule type: protein
 A:Residues: 1-10 <BMR>
 A:Note: the natural and synthetic hormones have the same biological activity
 C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
 F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 31.7%; Score 58; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.074;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 |||||
 Db 2 HWSYGLRPG 10

RESULT 8
 I78541
 gonadoliberin precursor - rhesus macaque (fragment)
 N:Alternate names: luteinizing hormone releasing hormone
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
 C:Accession: I78541
 R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
 Neuroendocrinology 60, 346-359, 1994
 A:Title: Developmental expression of the genes encoding transforming growth factor alpha
 A:Reference number: I58134; MUID:95124501
 A:Accession: I78541
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-67 <RES>
 A:Cross-references: GB:S75918; NID:g912831; PIDN:AA33096.1; PID:g912832
 C:Superfamily: gonadoliberin

Query Match 31.7%; Score 58; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.61;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 |||||
 Db 7 HWSYGLRPG 15

RESULT 9
 I51423
 gonadoliberin precursor - African clawed frog
 N:Alternate names: luteinizing hormone releasing hormone
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
 C:Accession: I51423
 R:Hayes, W.P.; Wray, S.; Battey, J.F.
 Endocrinology 134, 1835-1845, 1994
 A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved doma
 A:Reference number: I51423; MUID:94185563
 A:Accession: I51423
 A:Molecule type: DNA
 A:Title: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-89 <HAY>
 A:Cross-references: GB:I28040; NID:g496291; PIDN:AAA49728.1; PID:g496292
 C:Genetics:
 A:Gene: GnRH-I
 C:Superfamily: gonadoliberin

Query Match 31.7%; Score 58; DB 2; Length 89;
 Best Local Similarity 100.0%; Pred. No. 0.83;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 |||||
 Db 25 HWSYGLRPG 33

RESULT 10
 RHMSG
 gonadoliberin precursor - mouse
 N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releas
 N:Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
 C:Accession: A47578
 R:Masson, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli
 Science 234, 1366-1371, 1986
 A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible
 A:Reference number: A47578; MUID:87069928
 A:Accession: A47578
 A:Molecule type: DNA
 A:Residues: 1-90 <MAS>
 A:Cross-references: EMBL:M4872; NID:g193576; PIDN:AA37717.1; PID:g387175
 C:Genetics:
 A:Introns: 45/3; 77/3
 C:Function:
 A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitrop
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:22-31/Product: gonadoliberin #status predicted <GIB>
 F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
 F:22/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predic
 F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 31.7%; Score 58; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.84;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 |||||
 Db 23 HWSYGLRPG 31

RESULT 11
 RHMG
 gonadoliberin precursor [validated] - human
 N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releas
 N:Contains: gonadoliberin-associated protein (GAP); progadoliberin
 C:Species: Homo sapiens (man)
 C:Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
 C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
 R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
 Nucleic Acids Res. 17, 6403-6404, 1989
 A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone
 A:Reference number: S05308; MUID:89366682
 A:Accession: S05308
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-92 <HAY>
 A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
 R:Adelman, J.P.; Masson, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gona
 A:Reference number: A94090; MUID:86094338
 A:Accession: A26173
 A:Molecule type: mRNA
 A:Residues: 1-92 <ADE>
 A:Cross-references: GB:M12578; NID:g183418; PIDN:AA35916.1; PID:g386749
 A:Experimental source: hypothalamus

```

Mol. Biochem. Parasitol. 37, 275-280, 1989
A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell eg
A:Reference number: A60657; MUID:90114334
A:Accession: I60657
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 319-336,354-373 <LOC>
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:329-383/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match          49.7%; Score 91; DB 2; Length 405;
Best Local Similarity 55.3%; Pred. No. 0.0012;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAKERASSVFNVNS 36
Db 354 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNS 391

RESULT 3
OZQAF
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate IMT22)
C:Species: Plasmodium falciparum
C:Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
C:Accession: A03388
R:Date, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.
Science 225, 593-599, 1984
A:Title: Structure of the gene encoding the immunodominant surface antigen on the sporoz
A:Reference number: A03388; MUID:84250215
A:Accession: A03388
A:Molecule type: DNA
A:Residues: 1-412 <DAM>
A:Cross-references: GB:K02194; NID:g160160; PIDN:AAA29524.1; PID:g160161
A:Experimental source: clone 768
C:Comment: Residues 1-16 are the probable signal sequence.
C:Comment: There are 41 copies of a 4-residue repeating unit in the middle domain of the
F:336-390/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match          49.7%; Score 91; DB 1; Length 412;
Best Local Similarity 55.3%; Pred. No. 0.00013;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAKERASSVFNVNS 36
Db 361 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNS 398

RESULT 4
A54533
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thailand)
C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54533
R:del Portillo, H.A.; Nussenzweig, R.S.; Enea, V.
Mol. Biochem. Parasitol. 24, 289-294, 1987
A:Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
A:Reference number: A54533; MUID:87315205
A:Accession: A54533
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <DEL>
A:Cross-references: GB:M19752; NID:g160216; PIDN:AAA29555.1; PID:g160217
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:348-402/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match          49.7%; Score 91; DB 2; Length 424;
Best Local Similarity 55.3%; Pred. No. 0.00013;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAKERASSVFNVNS 36
Db 373 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNS 410

```

```

RESULT 5
A54529
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome)
C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54529
R:Lockyer, M.J.; Schwarz, R.T.
Mol. Biochem. Parasitol. 22, 101-108, 1987
A:Title: Strain variation in the circumsporozoite protein gene of Plasmodium falcipar
A:Reference number: A54529; MUID:87115616
A:Accession: A54529
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-442 <LOC>
A:Cross-references: GB:M15505; NID:g160214; PIDN:AAA29554.1; PID:g160215
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match          49.2%; Score 90; DB 2; Length 442;
Best Local Similarity 55.3%; Pred. No. 0.00019;
Matches 21; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAKERASSVFNVNS 36
Db 391 IKPGSADPKDQLDYENDIEKKICKMEKCSSVFNVNS 428

RESULT 6
RHPG
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAR>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A>Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A>Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match          31.7%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10

RESULT 7
RHSHG
gonadoliberin - sheep

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 8.38461 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHMSYGLRPGSSGSLDEKTIAMKRAKASVFNVNS 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.71:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	53.0	388	2 A39756	circumsporozoite p
2	91	49.7	405	2 S05428	circumsporozoite p
3	91	49.7	412	1 OZ2QAF	circumsporozoite p
4	91	49.7	424	2 A54533	circumsporozoite p
5	90	49.2	442	2 A54529	circumsporozoite p
6	58	31.7	10	1 RHPCG	gonadoliberin - pl
7	58	31.7	10	1 RHSHG	gonadoliberin - sh
8	58	31.7	67	2 I78541	gonadoliberin prec
9	58	31.7	89	2 I51423	gonadoliberin prec
10	58	31.7	90	1 RHMSG	gonadoliberin prec
11	58	31.7	92	1 RHUHG	gonadoliberin prec
12	58	31.7	92	1 RHRTG	gonadoliberin prec
13	58	31.7	332	1 OZ2QMB	circumsporozoite p
14	58	31.7	348	1 OZ2QMB	circumsporozoite p
15	55	30.1	98	2 I50739	gonadoliberin prec
16	55	30.1	264	2 A44969	circumsporozoite p
17	55	30.1	367	1 OZ2QMT	circumsporozoite p
18	54	29.5	10	1 RHAQI	gonadoliberin I -
19	54	29.5	92	2 I50644	gonadoliberin I pr
20	53.5	29.2	90	2 A23735	gonadoliberin prec
21	53	29.0	70	2 AC2866	hypothetical prote
22	53	29.0	501	2 T32848	hypothetical prote
23	52.5	28.7	487	2 A69645	methylnalonaate-sem
24	52	28.4	719	2 T52510	hypothetical prote
25	52	28.4	1401	2 T48079	hypothetical prote
26	51.5	28.1	80	1 RHIDS	gonadoliberin I pr
27	51.5	28.1	90	2 JC7395	salmon-type gonado
28	51.5	28.1	444	1 C64226	triglycer factor Mg2
29	51	27.9	315	2 F98295	hypothetical prote

30	51	27.9	332	2 A81667	conserved hypothet
31	51	27.9	332	2 B71508	hypothetical prote
32	51	27.9	429	2 A54504	circumsporozoite p
33	51	27.9	485	2 A60610	circumsporozoite p
34	51	27.9	959	2 S48962	MSH1 protein - yea
35	50.5	27.6	90	2 I51095	gonadoliberin prec
36	50.5	27.6	193	2 S43432	glutathione transf
37	50.5	27.6	482	2 B91069	succinate-semialde
38	50.5	27.6	482	2 D85913	succinate-semialde
39	50	27.3	223	1 A41031	glutathione transf
40	50	27.3	404	2 A64938	probable polygalac
41	50	27.3	433	2 B90288	thiamin biosynthes
42	50	27.3	444	2 C64498	probable polygalac
43	50	27.3	444	2 C64498	probable polygalac
44	50	27.3	590	2 D69722	thiamin biosynthes
45	50	27.3	595	2 E83891	thiamin biosynthes

ALIGNMENTS

RESULT 1

A39756 circumsporozoite protein - Plasmodium relictchenowi

C:Species: Plasmodium relictchenowi

C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999

C:Accession: A39756

R:Lai, A.A.; Goldman, I.F.

J. Biol. Chem. 266, 6686-6689, 1991

A>Title: Circumsporozoite protein gene from Plasmodium relictchenowi, a chimpanzee malar

A:Accession: A39756; MUID:91201303

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-388 <LAI>

A:Cross-references: GB:M60972; NID:g160228; PIDN:AAA29561.1; PID:g160229

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:312-366/Domain: thrombospondin type 1 repeat homology <THRL>

Query Match 53.0%; Score 97; DB 2; Length 388;
Best Local Similarity 57.9%; Pred. No. 1.8e+05;
Matches 22; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

OY 7 LRPSSGSPSLD-----EKKIAMKRAKASVFNVNS 36

Db 337 IKPGSAGKPKQDLDYENDLEKICKMEKRCSSVFNVNS 374

RESULT 2

S05428 circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)

C:Species: Plasmodium falciparum

C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000

C:Accession: S05428; A45527; I60657

R:Campbell, J.R.

Nucleic Acids Res. 17, 5854, 1989

A>Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate

A:Reference number: S05428; MUID:89345189

A:Accession: S05428

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-405 <CAW>

A:Cross-references: EMBL:X15363

R:Caepers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.

Mol. Biochem. Parasitol. 35, 185-190, 1989

A>Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate

A:Reference number: A45527; MUID:89364998

A:Accession: A45527

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <CAS>

A:Cross-references: GB:M22982; GB:J04650; NID:g160168; PIDN:AAA29527.1; PID:g160169

R:lockyer, M.D.; Marsh, K.; Newbold, C.I.


```

;      LENGTH: 27 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide
us-09-303-323-41

```

```

Query Match      67.5%; Score 108; DB 4; Length 27;
Best Local Similarity 71.4%; Pred. No. 4.2e-10;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

```

```

QY      3 LSEIKGVIYVHRLEGVGEPGLHWSYGLRP 30
      :|||||:|:| | |||||
Db      1 ISEIKGVIYVHKIEGTIGE--HWSYGLRP 26

```

```

Search completed: October 10, 2002, 16:14:01
Job time : 7.55769 secs

```

US-09-303-323-55

Query Match 68.8%; Score 110; DB 4; Length 31;

Best Local Similarity 78.6%; Pred. No. 2.4e-10;

Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30

Db 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30

RESULT 13

US-09-100-414B-41

Sequence 41, Application US/09100414B

Patent No. 6025468

GENERAL INFORMATION:

APPLICANT: Wang, Chang YI

TITLE OF INVENTION: NOVEL LHRH PEPTIDE

TITLE OF INVENTION: IMMUNOGENS

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC Windows

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/100,414B

FILING DATE: 20-JUNE-1998

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maria H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4157

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-751-6849

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 27 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-100-414B-41

Query Match 67.5%; Score 108; DB 3; Length 27;

Best Local Similarity 71.4%; Pred. No. 4.2e-10;

Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30

Db 1 LSEIKGVIVHRLGEGVPSLHWSYGLRP 26

RESULT 14

US-09-100-414B-47

Sequence 47, Application US/09100414B

Patent No. 6025468

GENERAL INFORMATION:

APPLICANT: Wang, Chang YI

TITLE OF INVENTION: NOVEL LHRH PEPTIDE

TITLE OF INVENTION: IMMUNOGENS

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC Windows

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/100,414B

FILING DATE: 20-JUNE-1998

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maria H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4157

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-751-6849

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 27 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-100-414B-47

Query Match 67.5%; Score 108; DB 3; Length 27;

Best Local Similarity 75.0%; Pred. No. 4.2e-10;

Matches 21; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30

Db 1 MSEIKGVIVHRLGEGVPSLHWSYGLRP 26

RESULT 15

US-09-303-323-41

Sequence 41, Application US/09303323

Patent No. 6228987

GENERAL INFORMATION:

APPLICANT: Wang, Chang YI

TITLE OF INVENTION: NOVEL LHRH PEPTIDE

TITLE OF INVENTION: IMMUNOGENS

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC Windows

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/303,323

FILING DATE: 30-APR-1999

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/100,414

FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maria H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4157

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-751-6849

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

Db 1 LSEIKGVIYHKLKLEGVGE--HWSYGLRP 26

RESULT 10
US-09-303-323-43
Sequence 43, Application US/09303323
Patent No. 6228987

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-09-303-323-43

Query Match 70.6%; Score 113; DB 4; Length 27;
Best Local Similarity 82.1%; Pred. No. 7,2e-11;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIYHKLKLEGVGPSLHWSYGLRP 30
|||||:|||||
Db 1 LSEIKGVIYHKLKLEGVGE--HWSYGLRP 26

RESULT 11
US-09-100-414B-55
Sequence 55, Application US/09100414B
Patent No. 6025468

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-09-100-414B-55

Query Match 68.8%; Score 110; DB 3; Length 31;
Best Local Similarity 78.6%; Pred. No. 2.4e-10;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LSEIKGVIYHKLKLEGVGPSLHWSYGLRP 30
|||||:|||||
Db 3 LSEIKGVIYHKLKLEGVLFGEHWSYGLRP 30

RESULT 12
US-09-303-323-55
Sequence 55, Application US/09303323
Patent No. 6228987

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/446,692
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-33

Query Match 72.5%; Score 116; DB 1; Length 45;
Best Local Similarity 85.7%; Pred. No. 4.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 3 LSEIKGVIVHRLGVGPELSHWSYGLRP 30
|||||
DB 19 LSEIKGVIVHRLGVGGE--HWSYGLRP 44

RESULT 8
US-08-488-351A-33
Sequence 33, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang YI
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/488,351A
APPLICATION NUMBER: US/08/488,351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-33

Query Match 72.5%; Score 116; DB 2; Length 45;
Best Local Similarity 85.7%; Pred. No. 4.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 3 LSEIKGVIVHRLGVGPELSHWSYGLRP 30
|||||
DB 19 LSEIKGVIVHRLGVGGE--HWSYGLRP 44

RESULT 9
US-09-100-414B-43
Sequence 43, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang YI
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA: US/09/100,414B
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-43

Query Match 70.6%; Score 113; DB 3; Length 27;
Best Local Similarity 82.1%; Pred. No. 7.2e-11;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

OY 3 LSEIKGVIVHRLGVGPELSHWSYGLRP 30

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-19

Query Match 72.5%; Score 116; DB 2; Length 27;
Best Local Similarity 85.7%; Pred. No. 2.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIYHRLEGVGPSSLMHSGLRP 30
|||||
Db 1 LSEIKGVIYHRLEGVGE--HWSYGLRP 26

RESULT 5
US-09-100-414B-36
Sequence 36, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-36

Query Match 72.5%; Score 116; DB 3; Length 27;

Best Local Similarity 85.7%; Pred. No. 2.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 3 LSEIKGVIYHRLEGVGPSSLMHSGLRP 30
|||||
Db 1 LSEIKGVIYHRLEGVGE--HWSYGLRP 26

RESULT 6
US-09-303-323-36
Sequence 36, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-36

Query Match 72.5%; Score 116; DB 4; Length 27;
Best Local Similarity 85.7%; Pred. No. 2.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIYHRLEGVGPSSLMHSGLRP 30
|||||
Db 1 LSEIKGVIYHRLEGVGE--HWSYGLRP 26

RESULT 7
US-08-446-692-33
Sequence 33, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin

```

RESULT 2
US-08-488-351A-20
: Sequence 20, Application US/08488351A
: Patent No. 5843346
: GENERAL INFORMATION:
: APPLICANT: Ladd, Anna
: APPLICANT: Wang, Chang YI
: APPLICANT: Zamb, Timothy
: TITLE OF INVENTION: Immunogenic LHRH peptide constructs
: TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
: NUMBER OF SEQUENCES: 114
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Maria C.H. Lin
: STREET: 345 Park Avenue
: CITY: New York
: STATE: NY
: COUNTRY: US
: ZIP: 10154-0053
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/488,351A
: FILING DATE: 7-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/446,692
: FILING DATE: 7-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/229,275
: FILING DATE: 14-APR-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/057,166
: FILING DATE: 27-APR-1992
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Maria C.H. Lin
: REGISTRATION NUMBER: 29,323
: REFERENCE/DOCKET NUMBER: 1151-4146 US2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212)415-8745
: TELEFAX: (516)751-6849
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 42 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-488-351A-20

Query Match 73.1%; Score 117; DB 2; Length 42;
Best Local Similarity 82.8%; Pred.No.2,9e-11;
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

OY 2 ILSEIKGVIHRLEGVEGSPSLHWSYGLRP 30
: ||||| ||||| ||||| ||||| |||||
Db 15 VLSEIKGVIHRLEGVGGE--HMSYGLRP 41

RESULT 3
US-08-446-692-19
: Sequence 19, Application US/08446692
: Patent No. 5759551
: GENERAL INFORMATION:
: APPLICANT: Ladd, Anna
: APPLICANT: Wang, Chang YI
: APPLICANT: Zamb, Timothy
: TITLE OF INVENTION: Immunogenic LHRH peptide constructs

```

```

      TITLE OF INVENTION: and synthetic universal immune stimulators for vaccine
      NUMBER OF SEQUENCES: 114
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Maria C.H. Lin
      STREET: 345 Park Avenue
      CITY: New York
      STATE: NY
      COUNTRY: US
      ZIP: 10154-0053
      COMPUTER READABLE FORM:
      MEDIUM TYPE: floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/446,692
      FILING DATE: 7-JUN-1995
      CLASSIFICATION: 424
      ATTORNEY/AGENT INFORMATION:
      NAME: Maria C.H. Lin
      REGISTRATION NUMBER: 29,323
      REFERENCE/DOCKET NUMBER: 1151-4146 US2
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212)415-8745
      TELEFAX: (516)751-6849
      INFORMATION FOR SEQ ID NO: 19:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 27 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: peptide
      US-08-446-692-19

Query Match          72.5%; Score 116; DB 1; Length 27;
Best Local Similarity 85.7%; Pred. No. 2,5e-11;
Matches    24; Conservative   0; Mismatches    2; Indels     2; Gaps     1;

QY       3 LSEIKGVIVHRLEGVGEPGLHWSYGILRP 30
         |||||||-----|||
Db        1 LSEIKGVIVHRLEGVGE--HWSYGILRP 26

RESULT 4
US-08-488-351A-19
: Sequence 19, Application US/08488351A
: Patent No. 5843446
: GENERAL INFORMATION:
: APPLICANT: Ladd, Anna
: APPLICANT: Wang, Chang Yi
: APPLICANT: Zamb, Timothy
: TITLE OF INVENTION: Immunogenic LHRH peptide constructs
: TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
: NUMBER OF SEQUENCES: 114
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Maria C.H. Lin
: STREET: 345 Park Avenue
: CITY: New York
: STATE: NY
: COUNTRY: US
: ZIP: 10154-0053
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/488,351A
: FILING DATE: 7-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/446,692
: FILING DATE: 7-JUN-1995
: CLASSIFICATION: 424

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 6.55769 Seconds

(Without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-9

Perfect score: 160

Sequence: 1 KLSEIKGVIVHRLGVGSPSLHMSGLRXP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	73.1	42	1	US-08-446-692-20
2	117	73.1	42	1	US-08-488-351A-20
3	116	72.5	27	1	US-08-446-692-19
4	116	72.5	27	2	US-08-488-351A-19
5	116	72.5	27	3	US-09-100-414B-36
6	116	72.5	27	4	US-09-303-323-36
7	116	72.5	27	1	US-08-446-692-33
8	116	72.5	45	2	US-08-488-351A-33
9	113	70.6	27	4	US-09-100-414B-43
10	113	70.6	27	4	US-09-303-323-43
11	110	68.8	31	3	US-09-100-414B-55
12	110	68.8	31	4	US-09-303-323-55
13	108	67.5	27	3	US-09-100-414B-41
14	108	67.5	27	3	US-09-100-414B-47
15	108	67.5	27	4	US-09-303-323-41
16	108	67.5	27	4	US-09-303-323-47
17	108	67.5	45	3	US-09-100-414B-45
18	107	66.9	31	4	US-09-303-323-45
19	107	66.9	31	4	US-09-100-414B-59
20	107	66.9	31	4	US-09-303-323-59
21	106	66.2	28	3	US-09-100-414B-38
22	106	66.2	28	4	US-09-303-323-38
23	105	65.6	31	4	US-09-100-414B-53
24	105	65.6	31	4	US-09-303-323-53
25	105	65.6	47	4	US-09-100-414B-60
26	105	65.6	47	4	US-09-303-323-60
27	105	65.6	49	3	US-09-100-414B-57

28	105	65.6	49	4	US-09-303-323-57	Sequence 57, Appl
29	103.5	64.7	25	1	US-08-446-692-17	Sequence 17, Appl
30	103.5	64.7	25	2	US-08-488-351A-17	Sequence 17, Appl
31	103	64.4	27	3	US-09-100-414B-50	Sequence 50, Appl
32	103	64.4	27	4	US-09-303-323-50	Sequence 50, Appl
33	103	64.4	35	3	US-09-100-414B-60	Sequence 80, Appl
34	103	64.4	35	4	US-09-303-323-80	Sequence 80, Appl
35	100	62.5	28	1	US-08-446-692-38	Sequence 38, Appl
36	100	62.5	28	2	US-08-488-351A-38	Sequence 38, Appl
37	100	62.5	46	1	US-08-446-692-40	Sequence 40, Appl
38	100	62.5	46	2	US-08-488-351A-40	Sequence 40, Appl
39	100	62.5	47	3	US-09-100-414B-63	Sequence 63, Appl
40	100	62.5	47	4	US-09-303-323-63	Sequence 63, Appl
41	99	61.9	28	3	US-09-100-414B-39	Sequence 39, Appl
42	99	61.9	28	4	US-09-303-323-39	Sequence 39, Appl
43	94	58.8	28	3	US-09-100-414B-37	Sequence 37, Appl
44	94	58.8	28	4	US-09-303-323-37	Sequence 37, Appl
45	87	54.4	34	5	PCR-US95-13841-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-08-446-692-20
; Sequence 20, Application US/08446692

; Patent No. 5759551

; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna

; APPLICANT: Zamb, Timothy

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

; TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin

; REGISTRATION NUMBER: 29, 323

; REFERENCE/DOCKET NUMBER: 1151-4146 US2

; TELEPHONE: (212)415-8745

; TELEFAX: (516)751-6849

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 42 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-446-692-20

Query Match 73.1%; Score 117; DB 1; Length 42;

Best Local Similarity 82.8%; Pred. No. 2.9e-11; Mismatches 2; Indels 2; Gaps 1;

Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

2 KLSEIKGVIVHRLGVGSPSLHMSGLRXP 30

15 VLSEIKGVIVHRLGVGGE--HWSYGLRXP 41

CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IGE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasion
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 31 AA;

Query Match 65.6%; Score 105; DB 21; Length 31;
 Best Local Similarity 67.9%; Pred. No. 7.4e-09;

Matches 19; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 3 LSEIKGVIVHRLEGVPSLHMSYGLRP 30
 :|||||||:|||||
 Db 3 LSEIKGVIVHRLEGVPSLHMSYGLRP 30

Search completed: October 10, 2002, 16:05:06
 Job time : 16.2949 secs

KM Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KM cholesterol ester transport protein; anti-arteriosclerotic.
 XX
 OS Chimeric - Measles virus.
 OS Chimeric - Rattus sp.
 XX
 PN W09966957-A2.
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13975.
 PR 20-JUN-1998; 98US-0100412.
 PA (UNBI-) UNITED BIOMEDICAL INC.
 PI Wang CY;
 DR WPI: 2000-160564/14.
 XX
 PT New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus -
 PS Example 1; Page 78; 129pp; English.
 XX
 CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVA Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IGE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVA Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVA Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive

CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 CC
 XX
 XX Sequence 28 AA:
 S0
 Query Match 66.2%; Score 106; DB 21; Length 28;
 Best Local Similarity 64.3%; Pred. No. 4.6e-05;
 Matches 18; Conservative 6; Mismatches 2; Indels 2; Gaps 1;
 QY 3 LSEIKGVYVHRLEGVGSPSLMWSIGLRP 30
 Db 2 ISEIKGIIIRHRIEGIGE--HMSYGLRP 27
 RESULT 15
 AAY91173
 ID AAY91173 standard; peptide: 31 AA.
 XX
 AC AAY91173;
 XX
 DT 22-MAY-2000 (first entry)
 XX
 DE Modified MVA Th epitope/LHRH antigenic peptide, SEQ ID NO:53.
 XX
 XX Promiscuous T-cell epitope; measles virus F protein; MVA;
 KM hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KM luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KM foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
 KM Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KM cholesterol ester transport protein; anti-arteriosclerotic.
 XX
 OS Chimeric - Measles virus.
 OS Chimeric - Rattus sp.
 XX
 PN W09966957-A2.
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13975.
 PR 20-JUN-1998; 98US-0100412.
 PA (UNBI-) UNITED BIOMEDICAL INC.
 PI Wang CY;
 DR WPI: 2000-160564/14.
 XX
 XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus -
 PS Example 1; Page 83; 129pp; English.
 XX
 CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target

CC Invasin domain immunostimulatory peptide of *Yersinia sp.*, a
CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
CC hormone-releasing hormone (LHRH). The synthetic Th epitope is derived
CC from a structured synthetic antigen library (SSAL) designated SSAL1 Th1.
CC SSAL1 Th1 is modelled after a promiscuous epitope taken from the F protein
CC of the Measles virus. The peptide immunogens cause induction of a
CC specific immune response to LHRH which is involved in regulation of
CC spermatogenesis, ovulation, oestrus, sexual development and secretion
CC of sex hormones. Provision of a promiscuous T helper epitope (which is
CC functional in genetically diverse subjects) provides optimum
CC immunogenicity to the B cell epitopes of the target antigen and thus
CC high antibody titres against the target antigen. The peptide immunogens
CC as (reversible) contraceptive: control of hormone-dependent tumours
CC (cancer of prostate or breast, also endometriosis); to prevent boar
CC taint (and improve meat quality) and for immunocastration.

SO Sequence 45 AA;

Query Match 67.5%; Score 108; DB 21; Length 45;
Best Local Similarity 71.4%; Pred. No. 4e-09;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

OY 3 LSEIKGVVHRLEGVSGPSLHMSYGLRP 30
Db 19 LSEIKGVVHRLEGVSGPSLHMSYGLRP 44

RESULT 13
AAV91179 standard; peptide: 31 AA.

XX AAV91179;
AC AAV91179;
XX 22-MAY-2000 (first entry)
DT 22-MAY-2000 (first entry)
DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:59.

XX Promiscuous T-cell epitope: measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
KW cholesterol ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.
OS Chimeric - Rattus sp.
XX W09966957-A2.
PN 29-DEC-1999.
PD 21-JUN-1999; 99WO-US13975.
XX 20-JUN-1998; 98US-0100412.
PR (UNBI-) UNITED BIOMEDICAL INC.
PA Wang CY;
PI WPI; 2000-160564/14.
DR WPI; 2000-160564/14.
XX New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus
XX Example 1; Page 86; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,

CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAV91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAV91122-V91142,
CC AAV91226 and AAV91245-V91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAV91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAV91144-V91155 are synthetic epitopes derived from this HBV epitope.
CC AAV91156-V91196, AAV91227 and AAV91242-V91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAV91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAV91200 is somatostatin, and AAV91201-V91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC peptides may be used to promote growth in livestock. AAV91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAV91209-V90211 are MVA Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAV90212 is a modified version of a human IGE
CC (immunoglobulin E) CH3 domain, and AAV90213-V90219 are Th epitopes/IGE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAV91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAV91221-V91222 comprise this peptide and a Th
CC epitope. AAV91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAV91224-V91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AAV91228-V91231 represent
CC CERP-derived peptides and AAV91232-V91241 are immunogens comprising a
CC CERP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAV91247 and AAV91252-V91257
CC are HIV-1 neutralising B-cell epitopes, and AAV91248-V91251 and
CC AAV91258-V91273 are antigenic peptides comprising MVA Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAV91198 and AAV91199 are respectively an immunostimulatory invasin
CC protein epitope from *Yersinia species*, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.

SO Sequence 31 AA;

Query Match 66.9%; Score 107; DB 21; Length 31;
Best Local Similarity 75.0%; Pred. No. 3.7e-09;
Matches 21; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 LSEIKGVVHRLEGVSGPSLHMSYGLRP 30
Db 3 LSEIKGVVHRLEGVSGPSLHMSYGLRP 30

RESULT 14
AAV91158
ID AAV91158 standard; peptide: 28 AA.

XX AAV91158;
AC AAV91158;
XX 22-MAY-2000 (first entry)
DT 22-MAY-2000 (first entry)
DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:38.

XX Promiscuous T-cell epitope: measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;

KM luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KM foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KM Plasmodium falciparum; circumsporozoite; anti-malarial; CERP;
 KM cholesterol ester transport protein; anti-arteriosclerotic.
 XX
 OS Chimeric - Measles virus.
 OS Chimeric - Rattus sp.
 XX
 PN W09966957-A2.
 XX
 PD 29-DEC-1999.
 XX
 XX 21-JUN-1999; 99WO-US13975.
 PF
 PR 20-JUN-1998; 98US-0100412.
 PA (UNB1-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 DR WPI; 2000-160564/14.
 PT New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus -
 XX
 PS Example 1; Page 80; 129pp; English.
 XX
 XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (WVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC WVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVA Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VPI capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVA Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and

CC AAY91258-Y91273 are antigenic peptides comprising MVA Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 XX
 XX
 SQ Sequence 45 AA;
 Query Match 67.5%; Score 108; DB 21; Length 45;
 Best Local Similarity 71.4%; Pred. No. 4e-09;
 Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
 Db 3 LSEIKGVIVHRLFCGVEGSPHSYGLRP 30
 :|||||||:|:|:|
 19 LSEIKGVIVHKEIGIGE--HWSYGLRP 44
 RESULT 12
 AAY68573
 ID AAY68573 standard; peptide; 45 AA.
 XX
 AC AAY68573;
 XX
 DT 05-MAY-2000 (first entry)
 XX
 DE Peptide immunogen comprising a Th epitope and LHRH target antigen.
 XX
 KM Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
 KM luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KM oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KM vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KM breast cancer; endometriosis; boar taint; meat quality;
 KM invasin domain; immunocastration.
 XX
 OS Synthetic.
 OS Yersinia sp.
 OS Measles virus.
 OS Unidentified.
 OS
 FH Key
 FT Peptide 1..16 Location/Qualifiers
 FT /note= "invasin domain AAY68565"
 FT Peptide 17..18
 FT /note= "spacer"
 FT Peptide 19..33
 FT /note= "helper Th epitope AAY68544"
 FT Peptide 34..35
 FT /note= "spacer"
 FT Peptide 36..45
 FT /note= "LHRH antigenic epitope AAY68566"
 XX
 PN W09966952-A1.
 XX
 PD 29-DEC-1999.
 XX
 XX 21-JUN-1999; 99WO-US13960.
 PF
 PR 20-JUN-1998; 98US-0100414.
 PA (UNB1-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 DR WPI; 2000-160562/14.
 PT New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer -
 XX
 PS Claim 9; Page 71; 102pp; English.
 CC The present sequence represents a peptide immunogen comprising an

PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
XX New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
XX Example 1; Page 84; 129pp; English.
PS

The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes and peptide immunogens along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response, specifically against Plasmodium falciparum, cholesteryl ester transport protein (CEPT) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration); for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AA91121 represents a promiscuous T helper epitope from the measles virus F (WVF) protein and sequences AA91122-91142, AA91226 and AA91245-91246 represent synthetic Th epitopes based on the WVF Th epitope. Sequence AA91143 represents a promiscuous Th epitope from Hepatitis B virus (HBV) surface antigen, and sequences AA91144-91155 are synthetic epitopes derived from this HBV epitope. AA91156-91196, AA91227 and AA91247-91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AA91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AA91200 is somatostatin, and AA91201-91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AA91208 is a human CD4 CDR2-like domain antigenic site, and AA91209-91201 are MWN Th epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. AA90212 is a modified version of a human IGE (immunoglobulin E) CH3 domain, and AA90213-910219 are Th epitope/IgE CH33 antigenic peptides which may be used in the treatment of allergies. AA91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AA91221-91222 comprise this peptide and a Th epitope. AA91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and AA91224-91225 comprise the CS antigen and an MWF Th epitope, and may be used in a malaria vaccine. AA91228-91231 represent CEHV-derived peptides and AA91232-91241 are Immunogens comprising a CEHV peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AA91247 and AA91252-91257 are HIV-1 neutralising B-cell epitopes, and AA91248-91251 and AA91258-91273 are antigenic peptides comprising MWN Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AA91198 and AA91199 are respectively an immunostimulatory invasive protein epitope from *Yersinia* species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

[illegible]

RESULT 9
AAV91161
ID AAV91161 standard; peptide; 27 AA.
AC AAV91161;
XX
XX 22-MAY-2000 (first entry)
DT
XX Modified MVE Th epitope/LHRH antigenic peptide, SEQ ID NO:41.
DE
XX
XX Promiscuous T-cell epitope; measles virus F protein; MVE;
KM hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KM luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; PMDV;
KM foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KM Plasmodium falciparum; circumsporozoite; antimalarial; CEMP;
KM cholesterol ester transport protein; anti-arteriosclerotic.
XX
XX Chimeric - Measles virus.
OS Chimeric - Rattus sp.
OS
XX WO966957-A2.
XX
XX 29-DEC-1999.
PD
XX 21-JUN-1999; 99MO-US13975.
XX
XX 20-JUN-1998; 98US-0100412.
PR
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
PA
XX Wang CY;
PI
XX WPI; 2000-160564/14.
XX
XX
XX New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
XX Example 1; Page 79; 129pp; English.
PS
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
XX and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CEMP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVE) protein and sequences AAY91122-Y91142,
CC AAY91126 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVE Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91247-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVN Th

CC		recurrent functional ovarian cysts, (severe) premenstrual syndrome or oestrogen-dependent breast cancer, or for induction of infertility.
XX		
XX		
SQ	Sequence	45 AA;
OY	Query Match	72.5%; Score 116; DB 15; Length 45;
	Best Local Similarity	85.7%; Pred. No. 2,5e-10;
	Matches	24; Conservative 0; Mismatches 2; Indels 2; Gaps
DB	3 LSEIKGVIVHRLGEGVPSLHMSTGLRP 30 19 LSEIKGVIVHRLGEGVF--HMSYGLRP 44 	
RESULT 7		
ID	AAV91163 standard; peptide: 27 AA.	
XX	AAV91163;	
AC	AAV91163;	
DT	22-MAY-2000 (first entry)	
XX		
DE	Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:43.	
XX		
KM	Promiscuous T-cell epitope: measles virus F protein; MVF;	
KM	hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;	
KW	luteinizing hormone releasing hormone; LHRH; contraceptive; anticancer;	
KW	somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; PMDV;	
KW	foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;	
KW	Plasmodium falciparum; circumsporozoite; antimalarial; CEMP;	
KX	cholesteryl ester transport protein; anti-arteriosclerotic.	
XX		
OS	Chimeric - Measles virus.	
OS	Chimeric - Rattus sp.	
XX		
PN	WO9966957-A2.	
PD	29-DEC-1999.	
XX		
XX	21-JUN-1999; 99WO-US13975.	
PF		
PR	20-JUN-1998; 98US-0100412.	
XX		
PA	(UNBI-) UNITED BIOMEDICAL INC.	
PI	Wang CY;	
DR	WPI: 2000-160564/14.	
PS		
PT	Example 1; Page 80; 129pp; English.	
XX		
CC	The invention relates to novel promiscuous T helper cell epitopes ('Th'),	
CC	and immunogenic peptides comprising the Th epitopes of the invention	
CC	along with B cell epitopes. The Th epitopes and peptide immunogens	
CC	containing them, are used to induce a T helper cell response,	
CC	specifically against Plasmodium falciparum, cholesteryl ester transport	
CC	protein (CEP) or HIV epitopes, but more generally against any pathogen,	
CC	immunoreactive self-antigen or tumour antigen. The Th epitopes and	
CC	peptide immunogens may be used for prevention and/or treatment of	
CC	infections (HIV, foot-and-mouth disease or malaria); for cancer	
CC	immunotherapy; for inhibition of the action of luteinising hormone	
CC	releasing hormone (LHRH) for contraception, treatment of hormone-	
CC	dependent cancer, prevention of boar taint in meat, and	
CC	immunocontraction; for promoting the growth of animals; or for	
CC	treating allergies or arteriosclerosis. Incorporation of a promiscuous	
CC	Th (functional in genetically diverse subjects) into an immunogen	
CC	improves capacity to induce a strong T helper cell-mediated immune	
CC	response, resulting in production of antibodies against a target	
CC	antigen. Th can replace carrier proteins and pathogen-derived T helper	

CC	epitopes. Sequence AAY91121 represents a promiscuous T helper epitope from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC	AAY91126 and AAY91245-Y91246 represent synthetic Th epitopes based on the MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC	from hepatitis B virus (HBV) surface antigen, and sequences
CC	AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC	AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC	is the LHRH target antigenic peptide used in these LHRH antigenic
CC	peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC	peptides comprising somatostatin and a Th epitope. Somatostatin
CC	immunogens may be used to promote growth in livestock. AAY91208 is a
CC	human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVN Th
CC	epitope/CyD4 CDR2 antigenic peptides which may be used to prevent HIV
CC	infection of T cells. AAY90212 is a modified version of a human IGE
CC	(immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC	antigenic peptides which may be used in the treatment of allergies.
CC	AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC	VPI capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC	epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC	antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC	epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC	CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC	CERP peptide and a Th epitope which may be used to prevent or treat
CC	arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC	are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC	AAY91258-Y91273 are antigenic peptides comprising MVN Th and HIV-1 B-cell
CC	epitope which may be used as a component in an anti-HIV-1 vaccine.
CC	AAY91198 and AAY91199 are respectively an immunostimulatory invasin
CC	protein epitope from Yersinia species, and hinge spacer peptide, both of
CC	which may optionally be used in the antigenic peptides of the
CC	invention.
SQ	Sequence 27 AA:
QY	Query Match 70.6%, Score 113, DB 21; Length 27;
	Best Local Similarity 82.1%; Pred. NO. 3.8e-10;
	Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
OY	3 LSEIKGVITVHRLEGVEGPELSHMSYGRP 30 1 LSEIKGVIVHKLEGGVGGE--IHMSYGLRP 26
Dd	
RESULT 8	
AAY91175	
ID	AAY91175 standard; peptide; 31 AA.
XX	
AC	AAY91175;
XX	
DT	22-MAY-2000 (first entry)
DE	Modified MVF Th epitope/LHRH antigenic peptide, SFO ID NO:55.
XX	
KW	Promiscuous T-cell epitope; measles virus F protein; MVF;
KW	hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW	luteinizing hormone releasing hormone; LHRH; contraceptive; anticancer;
KW	somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW	foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW	Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
KW	cholesteryl ester transport protein; anti-arteriosclerotic.
OS	Chimeric - Measles virus.
OS	Chimeric - Rattus sp.
XX	
PN	W09966957-A2.
PD	
XX	29-DEC-1999.
XX	
PF	21-JUN-1999; 99WO-US13975.
XX	
PR	20-JUN-1998; 98US-O100412.
XX	

PA (ZAMB/) ZAMB T.
 XX
 PI Ladd AE, Mang CY, Zamb T;
 XX
 DR WPI; 1994-357910/44.
 XX
 PT Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX
 PS Claim 8; Page 84; 213pp; English.
 XX
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 CC
 XX
 SQ Sequence 25 AA;
 Query Match 73.1%; Score 117; DB 15; Length 25;
 Best Local Similarity 85.7%; Pred. No. 8.6e-11;
 Matches 24; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 QY 3 LSEIKGVIHRLEGVGPGLHMSYGLRP 30
 :|||||
 Db 1 LSEIKGVIHRLEGVE---HMSYGLRP 24
 RESULT 2
 AAR62708
 ID AAR62708 standard; peptide; 42 AA.
 XX
 AC AAR62708;
 XX
 DT 10-SEP-1995 (first entry)
 XX
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; hapten;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Domain 1..15
 FT /note= "measles virus F protein helper T cell epitope"
 FT Domain 16..30
 FT /note= "measles virus F protein helper T cell epitope"
 FT Domain 33..42
 FT /note= "LHRH hapten"
 XX
 XX MO9425060-A.
 XX
 XX 10-NOV-1994.
 XX
 XX 28-APR-1994; 94WO-US04832.

PR 27-APR-1993; 93US-0057166.
 PR 14-APR-1994; 94US-0229275.
 XX
 XX
 PA (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX
 PI Ladd AE, Mang CY, Zamb T;
 XX
 DR WPI; 1994-357910/44.
 XX
 PT Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX
 PS Claim 8; Page 86; 213pp; English.
 XX
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 CC
 XX
 SQ Sequence 42 AA;
 Query Match 73.1%; Score 117; DB 15; Length 42;
 Best Local Similarity 82.8%; Pred. No. 1.6e-10;
 Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
 QY 2 LSEIKGVIHRLEGVGPGLHMSYGLRP 30
 :|||||
 Db 15 VLSEIKGVIHRLEGVGE--HMSYGLRP 41
 RESULT 3
 AAR62707
 ID AAR62707 standard; peptide; 27 AA.
 XX
 AC AAR62707;
 XX
 DT 10-SEP-1995 (first entry)
 XX
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; hapten;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Domain 1..15
 FT /note= "measles virus F protein helper T cell epitope"
 FT Domain 18..27
 FT /note= "LHRH hapten"
 XX
 XX MO9425060-A.
 XX
 XX 10-NOV-1994.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 16.2949 Seconds

(without alignments)
211.311 Million cell updates/sec

Title: US-09-848-834A-9

Perfect score: 160

Sequence: 1 KLSEIKGVIVHREGVGSPSLHMSYGLRPX 31

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:

747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_032802:*

1: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	73.1	25	AA62705	LHRH-containing im
2	117	73.1	42	AA62708	LHRH-containing im
3	116	72.5	27	AA62707	LHRH-containing im
4	116	72.5	27	AA62707	MVF Th epitope/LHR
5	116	72.5	27	AA62707	Peptide immunogen
6	116	72.5	45	AA62721	LHRH-containing im
7	113	70.6	27	AA62721	Modified MVF Th ep
8	110	68.8	31	AA62721	Modified MVF Th ep
9	108	67.5	27	AA62721	Modified MVF Th ep
10	108	67.5	27	AA62721	Modified MVF Th ep
11	108	67.5	45	AA62721	Modified MVF Th ep

12	108	67.5	45	21	AA62721	Peptide immunogen
13	107	66.9	31	21	AA62721	Modified MVF Th ep
14	106	66.2	28	21	AA62721	Modified MVF Th ep
15	105	65.6	31	21	AA62721	Modified MVF Th ep
16	105	65.6	31	21	AA62721	Modified MVF Th ep
17	105	65.6	47	21	AA62721	Modified MVF Th ep
18	105	65.6	47	21	AA62721	Modified MVF Th ep
19	105	65.6	49	21	AA62721	Modified MVF Th ep
20	103	64.4	27	21	AA62721	Modified MVF Th ep
21	103	64.4	27	21	AA62721	Modified MVF Th ep
22	103	64.4	35	21	AA62721	Modified MVF Th ep
23	100	62.5	28	15	AA62728	LHRH-containing im
24	100	62.5	46	15	AA62728	LHRH-containing im
25	100	62.5	47	21	AA62728	LHRH-containing im
26	100	62.5	47	21	AA62728	LHRH-containing im
27	99	61.9	28	21	AA62728	LHRH-containing im
28	94	58.8	28	21	AA62728	LHRH-containing im
29	89	55.6	27	16	AA62728	LHRH-containing im
30	87	54.4	34	17	AA62728	LHRH-containing im
31	84	52.5	27	21	AA62728	LHRH-containing im
32	84	52.5	27	21	AA62728	LHRH-containing im
33	84	52.5	45	21	AA62728	LHRH-containing im
34	84	52.5	45	21	AA62728	LHRH-containing im
35	81	50.6	31	21	AA62728	LHRH-containing im
36	81	50.6	31	21	AA62728	LHRH-containing im
37	79	49.4	27	21	AA62728	LHRH-containing im
38	79	49.4	27	21	AA62728	LHRH-containing im
39	79	49.4	27	21	AA62728	LHRH-containing im
40	79	49.4	27	21	AA62728	LHRH-containing im
41	79	49.4	45	21	AA62728	LHRH-containing im
42	78	48.8	40	20	AA62728	LHRH-containing im
43	78	48.8	40	20	AA62728	LHRH-containing im
44	78	48.8	40	20	AA62728	LHRH-containing im
45	78	48.8	40	21	AA62728	LHRH-containing im

ALIGNMENTS

RESULT 1

AA62705

AA62705 standard; peptide; 25 AA.

XX

AC AA62705;

XX

DT 10-SEP-1995 (first entry)

XX

DE LHRH-containing immunogenic peptide.

XX

KW Helper T cell epitope; universal immune stimulator; invasive; hapten;

KW vaccine; LHRH; interleukin hormone releasing hormone; prostate;

KW androgen-dependent carcinoma; antitumour; infertility;

KW measles virus F protein.

XX

OS Synthetic.

XX

EH Key

FT Domain

FT Domain

FT Domain

FT Domain

PN W09425060-A.

PN 10-NOV-1994.

PD 28-APR-1994; 94WO-US04832.

XX 27-APR-1993; 93US-0057166.

PR 14-APR-1994; 94US-0229275.

XX (LADD/) LADD A F.

PA (WANG/) WANG C Y.

Location/Qualifiers

1..15

/note="measles virus F protein helper T cell epitope"

16..25

/note="LHRH hapten"

DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WTF;
RX MEDLINE=99329215; PubMed=10400788;
RA Johnston I.C., Ter Meulen V., Schneider-Schaulles J.,
RA Schneider-Schaulles S.;
RT "A recombinant measles vaccine virus expressing wild-type
glycoproteins: consequences for viral spread and cell tropism.";
RL J. Virol. 73:6903-6915(1999).
DR EMBL; AJ133108; CAB38075.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59580 MW; 8255499968B5D862 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVYVHRLEGV 17
|||||
DB 288 LSEIKGVYVHRLEGV 302

RESULT 14
ID P90330 PRELIMINARY; PRT; 550 AA.
AC P90330;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGAHARA(HB);
RA Sheng J., Watanabe M., Ueda S.;
RT "Selection of a neurotropic variant of measles virus.";
RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGAHARA(HB);
RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced
fusogenicity of measles virus.";
RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; D63924; BAA09951.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59589 MW; 73E7BD457ABA39B7 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVYVHRLEGV 17
|||||
DB 288 LSEIKGVYVHRLEGV 302

RESULT 15
O93055

ID O93055 PRELIMINARY; PRT; 553 AA.
AC O93055;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE FUSION GLYCOPROTEIN PRECURSOR (FUSION PROTEIN).
GN F.
OS Measles virus, and
OS Measles virus (strain Edmonston-2agreb) (Subacute sclerose
OC panencephalitis virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234, 70149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HALLE;
RX MEDLINE=87224816; PubMed=3585281;
RA Buckland R., Gerald C., Barker R., Wild T.F.;
RT "Fusion glycoprotein of measles virus: nucleotide sequence of the gene
and comparison with other paramyxoviruses.";
RL J. Gen. Virol. 68:1695-1703(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VARIOUS STRAINS;
RA Parks C.L., Lerch R.A., Walpita P., Wang H.-P., Sidhu M.S., Udem S.A.;
RT "Comparison of predicted amino acid sequences from measles virus
strains in the Edmonston vaccine lineage.";
RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; D00090; BAA00056.1; -.
DR EMBL; AF266290; AAF85696.1; -.
DR EMBL; AF266288; AAF85680.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Signal.
FT SIGNAL. 1 26 POTENTIAL.
FT CHAIN 27 115 FUSION GLYCOPROTEIN F2 SUBUNIT.
FT CHAIN 116 553 FUSION GLYCOPROTEIN F1 SUBUNIT.
SQ SEQUENCE 553 AA; 59863 MW; 94C616494DDF8023 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVYVHRLEGV 17
|||||
DB 291 LSEIKGVYVHRLEGV 305

Search completed: October 10, 2002, 16:09:54
Job time : 14.0556 secs

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOYOSHIMA;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M., Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute sclerosing panencephalitis viruses: deduced amino acid sequences showed the cytoplasmic domain highly mutated --truncated, elongated or predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179432; AAF02697.1; -.
DR HSSP; P04849; 1SYE.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59504 MW; 2AA969D37FA5CA17 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIYHRLGV 17
|||||
DB 288 LSEIKGVIYHRLGV 302

RESULT 10

OQ0EW9 PRELIMINARY; PRT; 550 AA.
AC OQ0EW9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M., Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute sclerosing panencephalitis viruses: deduced amino acid sequences showed the cytoplasmic domain highly mutated --truncated, elongated or predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179436; AAF02701.1; -.
DR HSSP; P04849; 1SYE.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59405 MW; 0AE6DFC5DD22BBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIYHRLGV 17
|||||
DB 288 LSEIKGVIYHRLGV 302

RESULT 11

OQ0EW8 PRELIMINARY; PRT; 550 AA.
AC OQ0EW8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.

OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M., Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute sclerosing panencephalitis viruses: deduced amino acid sequences showed the cytoplasmic domain highly mutated --truncated, elongated or predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179437; AAF02702.1; -.
DR HSSP; P04849; 1SYE.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59315 MW; 086E51FED235EBBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIYHRLGV 17
|||||
DB 288 LSEIKGVIYHRLGV 302

RESULT 12

OQ0EW7 PRELIMINARY; PRT; 550 AA.
AC OQ0EW7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M., Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute sclerosing panencephalitis viruses: deduced amino acid sequences showed the cytoplasmic domain highly mutated --truncated, elongated or predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179438; AAF02703.1; -.
DR HSSP; P04849; 1SYE.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59353 MW; 086E51FED5582BBA CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIYHRLGV 17
|||||
DB 288 LSEIKGVIYHRLGV 302

RESULT 13

OQ0WK4 PRELIMINARY; PRT; 550 AA.
AC OQ0WK4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLEGV 17
|||||
DB 288 LSEIKGVIHRLEGV 302

RESULT 6

P90331 PRELIMINARY; PRT; 550 AA.
AC P90331;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGAHATA;
RA Sheng J., Watanabe M., Ueda S.;
RT Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGAHATA;
RA Sheng J., Watanabe M., Ueda S.;
RT Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.

RT "Selection of a neurotropic variant of measles virus."
RT "An amino acid alteration of F protein responsible for the enhanced
RT fusogenicity of measles virus."
RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGAHATA;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed."
RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL: D63926; BAA0958.1; -
DR EMBL: AF179431; AAF02696.1; -
DR HSSP: P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59530 MW; 97C991C7E2169839 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLEGV 17
|||||
DB 288 LSEIKGVIHRLEGV 302

RESULT 7

O9YJ94 PRELIMINARY; PRT; 550 AA.
AC O9YJ94;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.

OX NCBI_TaxID=11234;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9301V;
RX MEDLINE=98440529; PubMed=9765410;
RA Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,
RA Asakawa M., Nagai Y.;
RT "Measles virus attenuation associated with transcriptional impediment
RT and a few amino acid changes in the polymerase and accessory
RT proteins."
RL J. Virol. 72:8690-8696(1998).
DR EMBL: AB012949; BAA33877.1; -
DR EMBL: AB012948; BAA33871.1; -
DR HSSP: P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59512 MW; 7AA4F1D117197BF9 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLEGV 17
|||||
DB 288 LSEIKGVIHRLEGV 302

RESULT 8

O9OEX1 PRELIMINARY; PRT; 550 AA.
AC O9OEX1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MASUSAKO;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed."
RL Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF179430; AAF02695.1; -
DR HSSP: P04849; 1SVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59559 MW; 609EE024A7E59C54 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLEGV 17
|||||
DB 288 LSEIKGVIHRLEGV 302

RESULT 9

O9OEX0 PRELIMINARY; PRT; 550 AA.
AC O9OEX0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.

RESULT 2

004242 PRELIMINARY: PRT: 537 AA.

AC 004242: (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE FUSION PROTEIN.

GN F.

OS Measles virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.

OX NCBI_TaxID=11234;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86003063; PubMed=3167982;

RA Cattaneo R., Schmid A., Eschle D., Bacsko K., ter Meulen V.,

RA Billeter M.A.;

RT "Biased hypermutation and other genetic changes in defective measles
viruses in human brain infections.";

RL Cell 55:255-265(1988).

RN [2]

RP SEQUENCE FROM N.A.

RA Cattaneo R., Billeter M.A.;

RL Virology 0:0-0(0).

DR EMBL: X16567: CAA34574.1; -.

DR EMBL: X16567: CAA34575.1; -.

DR HSSP: P04849: 1SVF.

DR InterPro: IPR000776; Fusion_gly.

DR Pfam: PF00523; Fusion_gly; 1.

SO SEQUENCE 537 AA; 58275 MW; D0A60AC66D979E06 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 537;

Best Local Similarity 100.0%; Pred. No. 0.05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLGV 17

|||||

Db 291 LSEIKGVIHRLGV 305

RESULT 3

09PXA4 PRELIMINARY: PRT: 545 AA.

AC 09PXA4: (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE FUSION PROTEIN.

OS Measles virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.

OX NCBI_TaxID=11234;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-OSA-3;

RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,

RA Ogura H.;

RT "Nucleotide sequences of the fusion protein gene of subacute
sclerosing panencephalitis viruses: deduced amino acid sequences

RT showed the cytoplasmic domain highly mutated --truncated, elongated or

RT predicted secondary structure changed.";

RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

RL EMBL: AF119440; AAF02705.1; -.

RL EMBL: AF119440; AAF02705.1; -.

DR HSSP: P04849: 1SVF.

DR HSSP: P04849: 1SVF.

DR InterPro: IPR000776; Fusion_gly.

DR Pfam: PF00523; Fusion_gly; 1.

SO SEQUENCE 545 AA; 58907 MW; 0234C28AE193E77D CRC64;

Best Local Similarity 100.0%; Pred. No. 0.051;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLGV 17

|||||

Db 288 LSEIKGVIHRLGV 302

RESULT 4

091HA5

AC 091HA5: PRELIMINARY: PRT: 546 AA.

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DE FUSION PROTEIN.

GN F.

OS Rinderpest virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.

OX NCBI_TaxID=11241;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K;

RX MEDLINE=21014265; PubMed=11186456;

RA Alanot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,

RA Gusev A.A.;

RT "Primary structure of the F-gene from Rinderpest virus strain K.";

RL Mol. Gen. Microbiol. Virusol. 4:29-33(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-K;

RA Alanot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,

RA Gusev A.A.;

RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY035887; AAK63190.1; -.

DR EMBL: AY035887; AAK63190.1; -.

SO SEQUENCE 546 AA; 58572 MW; 449B2B2DD7405F08 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 546;

Best Local Similarity 100.0%; Pred. No. 0.051;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLGV 17

|||||

Db 284 LSEIKGVIHRLGV 298

RESULT 5

089495 PRELIMINARY: PRT: 550 AA.

AC 089495: (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE FUSION PROTEIN.

GN F.

OS Measles virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.

OX NCBI_TaxID=11234;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92230209; PubMed=1566568;

RA Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;

RT "Genetic variability of the glycoprotein genes of current wild-type

RT measles isolates.";

RL Virology 188:135-142(1992).

DR EMBL: M81903; AAA46422.1; -.

DR EMBL: M81901; AAA46421.1; -.

DR HSSP: P04849: 1SVF.

DR InterPro: IPR000776; Fusion_gly.

DR Pfam: PF00523; Fusion_gly; 1.

SO SEQUENCE 550 AA; 59564 MW; A78EC9CD6268E58 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 545;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 12.0556 Seconds

(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-9

Perfect score: 160

Sequence: 1 KLSLEIKGVIVHRLGEGVPSLIHMSYGLRXP 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	45.0	534	12	Q04243
2	72	45.0	537	12	Q04242
3	72	45.0	545	12	Q9PXA4
4	72	45.0	546	12	Q91HA5
5	72	45.0	550	12	Q89495
6	72	45.0	550	12	P90331
7	72	45.0	550	12	Q9YJ94
8	72	45.0	550	12	Q9QEX1
9	72	45.0	550	12	Q9QEX0
10	72	45.0	550	12	Q9QEW9
11	72	45.0	550	12	Q9QEW8
12	72	45.0	550	12	Q9QEW7
13	72	45.0	550	12	Q9QEW6
14	72	45.0	550	12	P90330
15	72	45.0	553	12	Q93055
16	72	45.0	553	12	Q83530

17	72	45.0	553	12	011383
18	72	45.0	553	12	Q83518
19	72	45.0	553	12	Q83521
20	72	45.0	553	12	Q83525
21	72	45.0	553	12	Q83527
22	72	45.0	553	12	Q83533
23	72	45.0	553	12	Q83536
24	72	45.0	553	12	Q91C36
25	72	45.0	553	12	Q91FK2
26	72	45.0	553	12	P88973
27	72	45.0	553	12	P88974
28	72	45.0	553	12	Q91248
29	72	45.0	553	12	Q04244
30	72	45.0	553	12	Q91OP2
31	72	45.0	553	12	Q9PW04
32	68	42.5	545	12	Q9QEW6
33	68	42.5	553	12	011380
34	66.5	41.6	552	12	Q66147
35	65	40.6	528	12	Q9JW9
36	65	40.6	662	12	Q9YKL7
37	65	40.6	662	12	Q89327
38	65	40.6	662	12	Q9DX22
39	65	40.6	662	12	Q91KN3
40	63.5	39.7	552	12	Q56852
41	63.5	39.7	552	12	Q66409
42	61	38.1	553	12	Q83629
43	60	37.5	546	12	Q84926
44	58	36.2	367	16	Q987M1
45	56	35.0	636	12	Q86486

ALIGNMENTS

RESULT 1
004243 PRELIMINARY: PRT: 534 AA.
AC 004243:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Bacsko K., ter Meulen V.,
RA Billeter M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RL Cell 55:255-265(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billeter M.A.;
RL Virology 0:0-0(0).
DR EMBL: X16568; CAA34581.1; -;
DR EMBL: X16568; CAA34582.1; -;
DR HSP: P04849; ISVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
SQ SEQUENCE 534 AA: 57899 MW: 637245E23B5E044 CRC64;

Query Match 45.0%; Score 72; DB 12; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGEGV 17
DB 291 LSEIKGVIVHRLGEGV 305

DR PIR; B36173; RHRTG.
 DR PIR; A48410; A48410.
 DR InterPro: IPR002012; GNRH.
 DR InterPro: IPR004079; Gonadoliberin1.
 DR Pfam: PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBERN1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 92
 FT PEPTIDE 24 92
 FT ACT_SITE 37 92
 FT ACT_SITE 26 26
 FT MOD_RES 24 24
 FT MOD_RES 33 33
 SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
 Query Match Best Local Similarity 35.6%; Score 57; DB 1; Length 92;
 Matches 12; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

OY 14 LEGVGGPSLHWSYGLRP 30
 DB 18 LEGCS--SQHWSYGLRP 32

RESULT 15

GONI_MACMU STANDARD; PRT; 67 AA.
 AC P55247;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Progonadoliberin I precursor [contains: Gonadoliberin I (LHRH I)
 DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I]
 DE (Fragment).
 GN GNRH1 OR GNRH OR LHRH.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecidae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RX MEDLINE=95124501; PubMed=7545971;
 RA Ma Y.J., Costa M.E., Ojeda S.R.;
 RT "Developmental expression of the genes encoding transforming growth
 RT factor alpha and its receptor in the hypothalamus of female rhesus
 RT macaques.";
 RL Neuroendocrinology 60:346-359(1994).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S75918; AAB33096.1; -
 CC InterPro: IPR002012; GNRH.
 CC Pfam; PF00446; GNRH; 1.
 CC PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;

KW Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 5
 FT CHAIN 6 >67
 FT PEPTIDE 6 15
 FT ACT_SITE 19 >67
 FT ACT_SITE 8 8
 FT MOD_RES 6 6
 FT MOD_RES 15 15
 FT MOD_RES 67 67
 FT NON_TER 67 67
 SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
 Query Match Best Local Similarity 33.8%; Score 54; DB 1; Length 67;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 21 SLHWSYGLRP 30
 DB 5 SQHWSYGLRP 14

Search completed: October 10, 2002, 16:06:28
 Job time : 6.0406 secs

FT CONFLICT 16 16 W -> S (IN REF. 3).
 SQ SEQUENCE 92 AA: 10380 MW: 30A72221B076FA79 CRC64;
 Query Match 35.9%; Score 57.5; DB 1; Length 92;
 Best Local Similarity 80.0%; Pred. No. 0.18;
 Matches 12; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 17 VEG-PSLHWSYGLRP 30
 ||| | |||||
 Db 18 VEGCSSQHWSTYGLRP 32

RESULT 13
 GONL_MOUSE STANDARD; PRT: 90 AA.
 AC P13562: 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 (luteinizing hormone releasing hormone I) (Gonadotropin releasing
 hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor
 I].
 GN GNRH1 OR GNRH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=87069928; PubMed=3024317;
 RA Macon A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
 RA Phillips H.S., Nikolics K., Seeburg P.H.;
 RT "A deletion truncating the gonadotropin-releasing hormone gene is
 responsible for hypogonadism in the hpg mouse.";
 RL Science 234:1366-1371(1986).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14872; AAA37717.1; -;
 DR MGD: MGI:95789; Gnrh.
 DR InterPro: IPR002012; GNRH.
 DR InterPro: IPR004079; GonadoliberinI.
 DR Pfam: PF00446; GNRH.1.
 DR PRINTS: PRO1541; GONADOLIBERI.
 DR PROSITE: PS00473; GNRH.1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 21
 FT CHAIN 90
 FT PEPTIDE 22 31 GONADOLIBERIN I.
 FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
 FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 90 AA: 10337 MW: 1C076FPA4826E4D9 CRC64;
 Query Match 35.6%; Score 57; DB 1; Length 90;
 Best Local Similarity 70.6%; Pred. No. 0.2;
 Matches 12; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

OY 14 LEGVEGPSLHWSYGLRP 30
 ||| | |||||
 Db 16 LEGCS--SQHWSTYGLRP 30

RESULT 14
 GONL_RAT STANDARD; PRT: 92 AA.
 AC P07490: 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 (luteinizing hormone releasing hormone I) (Gonadotropin releasing
 hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor
 I].
 GN GNRH1 OR GNRH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 of gonadotropin-releasing hormone and prolactin release-inhibiting
 factor in human and rat.";
 RT Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RL [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=89384661; PubMed=2476669;
 RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and
 hypothalamic expression.";
 RL Mol. Endocrinol. 3:1257-1262(1989).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Thymus;
 RX MEDLINE=93105480; PubMed=1468115;
 RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
 RT "Thymocytes express a mRNA that is identical to hypothalamic
 luteinizing hormone-releasing hormone mRNA.";
 RL Cell. Mol. Neurobiol. 12:447-454(1992).
 RN [4]
 RN SEQUENCE OF 1-47 FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=87149087; PubMed=3547652;
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
 RT "Two mammalian genes transcribed from opposite strands of the same
 DNA locus.";
 RL Science 235:1514-1517(1987).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S50870; AAB24572.1; -;
 DR EMBL: M12579; AAA41263.1; -;
 DR EMBL: M31670; AAA41264.1; -;
 DR EMBL: M15527; AAA42141.1; ALT_SEQ.
 DR EMBL: M15529; AAA42139.1; -;
 DR EMBL: M15528; -; NOT_ANNOTATED_CDS.

RP SEQUENCE OF 24-33.
 RX MEDLINE=72114303; PubMed=4946067;
 RA Baba Y., Matsuo H., Schally A.V.;
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.
 RT Confirmation of the proposed structure by conventional sequential
 RT analyses.";
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
 RN [3]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72065376; PubMed=4942726;
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 RT phase method.";
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
 RN [4]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72117544; PubMed=4946275;
 RA Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: L32864; AAA31066.1; -.
 CC PIR: A01411; RHPEG.
 DR InterPro: IPR002012; GNRH.
 DR InterPro: IPR004079; GonadoliberinI.
 DR Pfam: PF00446; GNRH.1.
 DR PRINTS: PR01541; GONADOLIBRNI.
 DR PROSITE: PS00473; GNRH.1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 23
 FT CHAIN 1 24 91 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SO SEQUENCE 91 AA; 10090 MW; 834047AF32DDAA99 CRC64;
 Query Match 35.9%; Score 57.5; DB 1; Length 91;
 Best Local Similarity 50.0%; Pred. No. 0.17;
 Matches 15; Conservative 4; Mismatches 8; Indels 3; Gaps 1;
 Oy 1 KLSEIKGYIVHREGEPSLHMSYGLRP 30
 Db 6 KLLALTLCTVCGSSQHSYGLRP 32
 RESULT 12
 GONI_HUMAN STANDARD; PRT; 92 AA.
 AC P01148;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 DE (luteinizing releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated
 DE peptide I].

GN GNRH1 OR GNRH OR LHRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89366682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 RT hormone gene.";
 RL Nucleic Acids Res. 17:6403-6403(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85012739; PubMed=6090951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 RT releasing hormone.";
 RL Nature 311:666-668(1984).
 RN [4]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=83126573; PubMed=6760865;
 RA Tan L., Rousseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide
 RT biosynthesized in the human placenta.";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 CC Lutrepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm
 CC (Sérono).
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X01059; CAA25526.1; -.
 DR EMBL: M12578; AAA35916.1; -.
 DR PIR: X15215; CAA33285.1; -.
 DR PIR: A01410; RHHUG.
 DR PIR: A26173; A26173.
 DR PIR: S05308; S05308.
 DR MIM: 152760; -.
 DR InterPro: IPR002012; GNRH.
 DR InterPro: IPR004079; GonadoliberinI.
 DR Pfam: PF00446; GNRH.1.
 DR PRINTS: PR01541; GONADOLIBRNI.
 DR PROSITE: PS00473; GNRH.1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pharmaceutical; Signal.
 FT SIGNAL 1 23
 FT CHAIN 1 24 92 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 FT

```

CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X16566; CA34567.1; -
DR      HSSP; P04849; ISVF.
DR      InterPro: IPR000776; Fusion_gly.
DR      Pfam; PF00523; fusion_gly.1.
KW      Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT      SIGNAL 1 26
FT      CHAIN 27 529
FT      CHAIN 27 115
FT      CHAIN 116 529
FT      TRANSMEM 116 139
FT      DOMAIN 140 497
FT      DOMAIN 498 518
FT      DOMAIN 519 529
FT      DISULFID 71 198
FT      CARBOHYD 32 64
FT      CARBOHYD 64 70
FT      CARBOHYD 70 70
SO      SEQUENCE 529 AA; 57331 MW; AE987BC9F07E9AA9 CRC64;

Query Match 40.0%; Score 64; DB 1; Length 529;
Best Local Similarity 93.3%; Pred. No. 0.12;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      3 LSEIKGVVHRLEGV 17
DB      291 LSEIKGVVHRLEGV 305

RESULT 10
VGLF_PHODV STANDARD; PRT; 631 AA.
AC      P28866;
DC      01-DEC-1992 (Rel. 24, Created)
DT      01-DEC-1992 (Rel. 24, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Fusion glycoprotein precursor [contains: Fusion glycoprotein F2;
DE      Fusion glycoprotein F1].
GN      F.
OS      Phocine distemper virus (PDV).
OC      Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX      NCBI_TaxID=11240;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-ISOLATE DK88-4A;
RX      MEDLINE=92113538; PubMed=1765768;
RA      Koevanees J., Blixenkron-Moeller M., Sharma B., Oerwell C.,
RA      Norby E.;
RT      "The nucleotide sequence and deduced amino acid composition of the
RT      haemagglutinin and fusion proteins of the morbillivirus Phocid
RT      distemper virus.";
RL      J. Gen. Virol. 72:2959-2966(1991).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-UISTER/88;
RX      MEDLINE=92398437; PubMed=1524494;
RA      Curran M.D., Lu Y.J., Rima B.K.;
RT      "The fusion protein gene of phocine distemper virus: nucleotide and
RT      deduced amino acid sequences and a comparison of morbillivirus fusion
RT      proteins.";
RL      Arch. Virol. 126:159-169(1992).
RN      [3]

```

```

RP      SEQUENCE OF 95-631 FROM N.A.
RC      STRAIN-UISTER/88;
RX      MEDLINE=91089508; PubMed=2264246;
RA      Curran M.D., Ioan D.O., Rima B.K., Kennedy S.;
RT      "Nucleotide sequence analysis of phocine distemper virus reveals its
RT      distinctness from canine distemper virus.";
RL      Vet. Rec. 127:430-431(1990).
CC      -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC      MEMBRANES.
CC      -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC      LINKED BY A DISULFIDE BOND.
CC      -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D10371; BA01206.1; -
DR      PIR; JQ1368; VGNZPD.
DR      PIR; A48346; A48346.
DR      HSSP; P04849; ISVF.
DR      InterPro: IPR000776; Fusion_gly.
DR      Pfam; PF00523; fusion_gly.1.
KW      Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT      SIGNAL 1 631
FT      CHAIN ? 631
FT      CHAIN ? 188
FT      CHAIN 194 631
FT      DISULFID 149 276
FT      TRANSMEM 89 106
FT      TRANSMEM 194 212
FT      TRANSMEM 575 595
FT      TRANSMEM 110 110
FT      CARBOHYD 142 142
FT      CARBOHYD 148 148
FT      CONFLICT 63 63
SO      SEQUENCE 631 AA; 68873 MW; D1FC87CDD42E69B8 CRC64;

Query Match 40.0%; Score 64; DB 1; Length 631;
Best Local Similarity 80.0%; Pred. No. 0.14;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      3 LSEIKGVVHRLEGV 17
DB      369 LSEIKGVVHRLEAV 383

RESULT 11
GONL_PIG STANDARD; PRT; 91 AA.
AC      P49921;
DC      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DE      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Progonaoliberin I precursor [contains: Gonadoliblerin I (LHRH I)
DE      (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE      hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].
GN      GNRH1 OR GNRH.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9623;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE-Hypothalamus;
RA      Weesner G.D., Mattern R.L., Becker B.A.;
RT      Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN      [2]

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M21514; AAA47400.1; -
CC PIR: A31051; VGNZRK.
CC HSSP: P04849; ISVF.
CC InterPro: IPR000776; Fusion_gly.
CC Pfam: PF00523; fusion_gly; 1.
CC Glycoprotein: Fusion protein; Transmembrane; Envelope protein; Signal.
CC SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFD 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58662 MW; 476D74DC18BCFCF CRC64;

Query Match 40.6%; Score 65; DB 1; Length 546;
Best Local Similarity 86.7%; Pred. No. 0.087; Indels 0; Gaps 0;
Matches 13; Conservative 1; Mismatches 1;

OY 3 LSEIKGVIHRLEGV 17
DB 284 LSEIKGVIHRLEGV 298
|||||:|:|:|
RESULT 8
VGLF_CDVO STANDARD; PRT; 662 AA.
AC P12569; O65991;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OC NCBI_TaxID=11233;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88129050; PubMed=3433924;
RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;
RT "The nucleotide sequence of the gene encoding the F protein of canine
RT distemper virus: a comparison of the deduced amino acid sequence with
RT other paramyxoviruses";
RT Virus Res. 8:373-386(1987).
RL [12]
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227696; PubMed=8470428;
RA Wild T.F., Bernard A., Spehner D., Villevet D., Drilling R.;
RT "Vaccination of mice against canine distemper virus-induced
RT encephalitis with vaccinia virus recombinants encoding measles or
RT canine distemper virus antigens";
RT Vaccine 11:438-444(1993).
RL [1]
RN [1]
RP FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN

CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M21849; AAA42878.1; -
CC PIR: J50321; VGNZCD.
CC HSSP: P04849; ISVF.
CC InterPro: IPR000776; Fusion_gly.
CC Pfam: PF00523; fusion_gly; 1.
CC Glycoprotein: Fusion protein; Transmembrane; Envelope protein; Signal.
CC SIGNAL 1 2
FT CHAIN 2 662 FUSION GLYCOPROTEIN F0.
FT CHAIN 2 224 PROTEIN F2.
FT CHAIN 225 662 PROTEIN F1.
FT TRANSMEM 606 629 POTENTIAL.
FT DISULFD 180 307 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 3 3 R -> K (IN REF. 2).
FT CONFLICT 140 140 D -> N (IN REF. 2).
FT CONFLICT 152 152 N -> S (IN REF. 2).
FT CONFLICT 171 171 I -> M (IN REF. 2).
FT CONFLICT 174 174 A -> V (IN REF. 2).
FT CONFLICT 662 662 L -> H (IN REF. 2).
SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;

Query Match 40.6%; Score 65; DB 1; Length 662;
Best Local Similarity 86.7%; Pred. No. 0.11; Indels 0; Gaps 0;
Matches 13; Conservative 1; Mismatches 1;

OY 3 LSEIKGVIHRLEGV 17
DB 400 LSEIKGVIHRLEGV 414
|||||:|:|:|
RESULT 9
VGLF_MEASI STANDARD; PRT; 529 AA.
AC P26031; O83298;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain IP-3-Ca) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
OC NCBI_TaxID=11237;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92263801; PubMed=1585658;
RA Schmidt A., Spielhofer P., Cattaneo R., Bacsko K., Ter Meulen V.,
RA Billeter M.A.;
RT "Subacute sclerosing panencephalitis is typically characterized by
RT alterations in the fusion protein cytoplasmic domain of the
RT persisting measles virus";
RT Virology 188:910-915(1992).
RL [1]
RN [1]
RP FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN


```
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088609; PubMed=7996154;
RA Evans S.A., Barton M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
RL J. Gen. Virol. 75:3611-3617(1994).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z30700; CA83186.1; -
DR EMBL; Z30697; CA83181.1; -
DR PIR; S47305; S47305.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly.1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58705 MW; ED3DF8AFDECB95 CRC64;

Query Match 44.4%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.012; Mismatches 0; Gaps 0;
Matches 14; Conservative 1; Indels 0;

QY 3 LSEIKGVTHRLGV 17
DB 284 LSEIKGVTHRLGV 298

RESULT 6
VGLF_RINDL STANDARD; PRT; 546 AA.
ID VGLF_RINDL
AC P10864;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain 1) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11243;
RN SEQUENCE FROM N.A.
RP MEDLINE=88219541; PubMed=3285575;
RA Tsukiyama K., Yoshikawa Y., Yamaguchi K.;
RT "Fusion glycoprotein (F) of rinderpest virus: entire nucleotide
RT sequence of the F mRNA, and several features of the F protein.";
```

```
RL Virology 164:523-530(1988).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20870; AAA47399.1; -
DR PIR; A28921; VGNZRL.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly.1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58911 MW; 985029418F28FFB5 CRC64;

Query Match 41.2%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.063; Mismatches 1; Gaps 0;
Matches 14; Conservative 0; Indels 1;

QY 3 LSEIKGVTHRLGV 17
DB 284 LSEIKGVTHRLGV 298

RESULT 7
VGLF_RINDK STANDARD; PRT; 546 AA.
ID VGLF_RINDK
AC P12574;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain Kabete O) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11242;
RN SEQUENCE FROM N.A.
RP MEDLINE=88322864; PubMed=3413983;
RA Hsu D., Yamanaka M., Miller J., Dale B., Grubman M., Yilma T.;
RT "Cloning of the fusion gene of rinderpest virus: comparative sequence
RT analysis with other morbilliviruses."
RL Virology 166:149-153(1988).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
```

DT 01-AUG-1988 (Rel. 08, Last Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
Fusion glycoprotein F1].
GN F.
OS Measles virus (strain Edmonston) (Subacute sclerose panencephalitis
virus).
OS Measles virus (strain Halle) (Subacute sclerose panencephalitis
virus).
OS Measles virus (strain Leningrad-16) (Subacute sclerose panencephalitis
virus).
OS Measles virus (strain Edmonston-Zagreb) (Subacute sclerose
panencephalitis virus).
OS Measles virus (strain Philadelphia-26) (Subacute sclerose
panencephalitis virus).
OS Measles virus (strain Edmonston B) (Subacute sclerose panencephalitis
virus).
OS Measles virus (strain Edmonston B) (Subacute sclerose panencephalitis
virus).
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=11235, 11236, 70147, 70149, 70148, 70146;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-EDMONSTON;
RX MEDLINE=67071668; PubMed=3788062;
RA Richardson C.D., Hull D., Greer P., Hasel K., Berkovich A.,
Englund G., Bellini W.J., Rima B., Lazzarini R.A.;
RT "The nucleotide sequence of the mRNA encoding the fusion protein of
measles virus (Edmonston strain): a comparison of fusion proteins
from several different paramyxoviruses.";
RL Virology 155:508-523(1986).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-HALLE;
RX MEDLINE=87224816; PubMed=3585281;
RA Buckland R., Gerald C., Barker R., Wild T.F.;
RT "Fusion glycoprotein of measles virus: nucleotide sequence of the
gene and comparison with other paramyxoviruses.";
RL J. Gen. Virol. 68:1695-1703(1987).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-EDMONSTON;
RX MEDLINE=90085790; PubMed=2596022;
RA Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Bacsko K.,
Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;
RT "Mutated and hypermutated genes of persistent measles viruses which
caused lethal human brain diseases.";
RL Virology 173:415-425(1989).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN-EDMONSTON;
RX MEDLINE=92265801; PubMed=1585658;
RA Schmid A., Spielhofer P., Cattaneo R., Bacsko K., Ter Meulen V.,
Billeter M.A.;
RT "Subacute sclerosing panencephalitis is typically characterized by
alterations in the fusion protein cytoplasmic domain of the
persisting measles virus.";
RL Virology 188:910-915(1992).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN-EDMONSTON, Leningrad-16, AND EDMONSTON-ZAGREB;
RX MEDLINE=94249283; PubMed=8191786;
RA Rota J.S., Wang Z.D., Rota P.A., Bellini W.J.;
RT "Comparison of sequences of the H, F, and N coding genes of measles
virus vaccine strains.";
RL Virus Res. 31:317-330(1994).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN-PHILADELPHIA-26;
RX MEDLINE=94303181; PubMed=8030232;
RA Hummel K.B., Vanchiere J.A., Bellini W.J.;
RT "Restriction of fusion protein mRNA as a mechanism of measles virus
persistence.";

RL Virology 202:665-672(1994).
[7]
RP SEQUENCE FROM N.A.
RC STRAIN-EDMONSTON B;
RA Billeter M.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
MEMBRANES.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14915; AAA46423.1; -!
DR EMBL: X05597; CAA29090.1; ALT_INIT.
DR EMBL: K01711; AAA75498.1; ALT_INIT.
DR EMBL: K01711; AAA75499.1; -!
DR EMBL: U03657; AAA56647.1; ALT_INIT.
DR EMBL: U03659; AAA56649.1; ALT_INIT.
DR EMBL: U03670; AAA56660.1; ALT_INIT.
DR EMBL: U08416; AAA50550.1; ALT_INIT.
DR EMBL: Z66517; CAA91367.1; ALT_INIT.
DR EMBL: Z66517; CAA91368.1; -!
DR PIR: A26962; VGNZMV.
DR HSSP: P04849; ISVF.
DR InterPro: IPR000776; Fusion_gly.
DR Pfam: PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 550 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT DISULFD 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA: 59532 MW: 7AA4F1CA82169093 CRC64;
Query Match 45.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 LSEIKGVIVHREGV 17
DB 288 LSEIKGVIVHREGV 302
RESULT 5
VQLE_RINDR 45.0%; PRT: 546 AA.
ID VQLE_RINDR STANDARD: PRT: 546 AA.
AC P41336;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain RBOK) (RDV).
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OC NCBI_TaxID=36409;

SO SEQUENCE 534 AA; 57963 MW; F5B21757E643844D CRC64;
Query Match 45.0%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LSEIKGVIHRLGV 17
|||||
DB 288 LSEIKGVIHRLGV 302
RESULT 2
VGLF_RINDB STANDARD; PRT; 546 AA.
AC P41360;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
Fusion glycoprotein F1].
F.
GN F.
OS Rinderpest virus (strain RBT1) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_FaxID=39007;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088609; PubMed=7996154;
RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus.";
RL J. Gen. Virol. 75:3611-3617(1994).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z31656; CAAB3482.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 O-LINKED (POTENTIAL).
SO SEQUENCE 546 AA; 58418 MW; 38B539B89344F401 CRC64;
Query Match 45.0%; Score 72; DB 1; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LSEIKGVIHRLGV 17
|||||

DB 284 LSEIKGVIHRLGV 298
RESULT 3
VGLF_MEASA STANDARD; PRT; 550 AA.
AC P35973;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
Fusion glycoprotein F1].
F.
GN F.
OS Measles virus (strain Aik-C) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_FaxID=36408;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227570; PubMed=8470368;
RA Mori T., Sasaki K., Hashimoto H., Makino S.;
RT "Molecular cloning and complete nucleotide sequence of genomic RNA of
RT the Aik-C strain of attenuated measles virus.";
RL Virus Genes 7:67-81(1993).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S58435; AAB26145.1; -.
DR PIR; E48556; E48556.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
FT CHAIN 24 112 PROTEIN F2.
FT CHAIN 113 550 PROTEIN F1.
FT TRANSMEM 113 136 POTENTIAL.
FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 495 515 POTENTIAL.
FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 550 AA; 59540 MW; AAC4DAB92DED0938 CRC64;
Query Match 45.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 LSEIKGVIHRLGV 17
|||||
DB 288 LSEIKGVIHRLGV 302
RESULT 4
VGLF_MEASE STANDARD; PRT; 550 AA.
AC P08300;
ID VGLF_MEASE

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 4.0406 Seconds

(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834a-9

Perfect score: 160

Sequence: 1 KLSEIKGVIVHLEGVGPELSLHWSYGLRPX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	45.0	534	1	VGIF_MEASY
2	72	45.0	546	1	VGIF_MEASY
3	72	45.0	550	1	VGIF_MEASA
4	72	45.0	550	1	VGIF_MEASE
5	71	44.4	546	1	VGIF_MEASE
6	66	41.2	546	1	VGIF_MEASE
7	65	40.6	546	1	VGIF_MEASE
8	65	40.6	546	1	VGIF_MEASE
9	64	40.0	529	1	VGIF_MEAST
10	64	40.0	631	1	VGIF_MEAST
11	57.5	35.9	91	1	VGIF_MEAST
12	57.5	35.9	92	1	VGIF_MEAST
13	57	35.6	90	1	VGIF_MEAST
14	57	35.6	92	1	VGIF_MEAST
15	54	33.8	67	1	VGIF_MEAST
16	53	33.1	508	1	VGIF_MEAST
17	52	32.5	61	1	VGIF_MEAST
18	52	32.5	63	1	VGIF_MEAST
19	52	32.5	89	1	VGIF_MEAST
20	52	32.5	92	1	VGIF_MEAST
21	52	32.5	379	1	VGIF_MEAST
22	50.5	31.6	92	1	VGIF_MEAST
23	50.5	31.6	99	1	VGIF_MEAST
24	49	30.6	582	1	VGIF_MEAST
25	48.5	30.3	95	1	VGIF_MEAST
26	48	30.0	10	1	VGIF_MEAST
27	48	30.0	92	1	VGIF_MEAST
28	48	30.0	94	1	VGIF_MEAST
29	48	30.0	95	1	VGIF_MEAST
30	48	30.0	95	1	VGIF_MEAST
31	48	30.0	124	1	VGIF_MEAST
32	47	29.4	122	1	VGIF_MEAST
33	47	29.4	452	1	VGIF_MEAST

34	47	29.4	452	1	HEMN_RHOSH	P95651 rhodobacter
35	46.5	29.1	207	1	TER2_ECOLI	P04483 escherichia
36	46.5	29.1	417	1	RF1_PYPAB	Q9V151 pyrococcus
37	46	28.7	110	1	YHB1_ACTAC	P96769 actinobacil
38	46	28.7	555	1	VGIF_P11HC	P12605 human parai
39	45.5	28.4	338	1	DCUP_AQUAE	O66667 aquifex aeo
40	45.5	28.4	417	1	RF1_PYPAB	O59264 pyrococcus
41	45	28.1	191	1	MOBA_PYPAB	O9V060 pyrococcus
42	45	28.1	235	1	Y381_MYCPN	P75219 mycoplasma
43	45	28.1	584	1	GUIND_CIOCE	P25472 clostridium
44	45	28.1	880	1	TYO3_MOUSE	P55144 mus musculi
45	45	28.1	1042	1	TLRH_METJA	Q60295 methanococc

ALIGNMENTS

RESULT 1	ID	VGIF_MEASY	STANDARD:	PRT:	534 AA.
AC	P26032:				
DT	01-MAY-1992 (Rel. 22, Created)				
DT	01-MAY-1992 (Rel. 22, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2; Fusion glycoprotein F1].				
GN	F.				
OS	Measles virus (strain Yamagata-1) (Subacute sclerosing panencephalitis virus).				
OC	Viruses: ssRNA negative-strand viruses; Mononegavirales;				
OC	Paramyxoviridae; Paramyxovirinae; Morbilliviruses.				
OX	NCBI:taxid=11239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90385702; PubMed=1698327;				
RA	Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;				
RT	"Molecular analysis of structural protein genes of the Yamagata-1 strain of defective subacute sclerosing panencephalitis virus. IV. Nucleotide sequence of the fusion gene."				
RT	Viruses 4:1173-181(1990).				
RL	FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR MEMBRANES.				
CC	- FUNCTION: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2 LINKED BY A DISULFIDE BOND.				
CC	- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN FAMILY.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	EMBL: D10548; BA001405.1; -.				
DR	PIR: J00274; J00274.				
DR	HSSP: P04849; ISVF.				
DR	InterPro: IPR000776; Fusion_gly.				
DR	Pfam: PF00523; fusion_gly; 1.				
KW	Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.				
FT	SIGNAL	1	23		
FT	CHAIN	24	534		
FT	CHAIN	24	112		
FT	CHAIN	113	534		
FT	TRANSMEM	113	136		
FT	TRANSMEM	137	494		
FT	DOMAIN	495	515		
FT	DOMAIN	516	534		
FT	DISULFID	68	195		
FT	CARBOHYD	29	29		
FT	CARBOHYD	61	61		
FT	CARBOHYD	67	67		

C:Superfamily: parainfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein
 F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
 F:89-106/Domain: transmembrane #status predicted <TM1>
 F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
 F:194-219/Domain: transmembrane #status predicted <TM2>
 F:575-595/Domain: transmembrane #status predicted <TM3>
 F:110,142/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 40.0%; Score 64; DB 1; Length 631;
 Best Local Similarity 80.0%; Pred. No. 0.51;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 LSEIKGVYHRLQGV 17
 |||:|||||
 Db 369 LSEVKGVVHRLQAV 383

Search completed: October 10, 2002, 16:12:02
 Job time : 9.29701 secs

C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 24-Nov-1999
C:Accession: J02223
R:Visser, I.K.G.; van der Heijden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oerwell
J. gen. Virol. 74, 1989-1994, 1993
A:Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites an
e virus entity.
A:Reference number: J02223; MUID:93389459
A:Accession: J02223
A:Molecule type: mRNA
A:Residues: 1-542 <VTS>
A:Cross-references: GB:L07075
A:Note: the authors translated the codon ATC for residue 4 as Leu
C:Comment: This fusion protein F0 is cleaved into F1 and F2 chains.
C:Genetics:
A:Gene: F
C:Superfamily: paramfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-15/Domain: signal sequence #status predicted <IG>
F:16-542/Product: fusion protein #status predicted <MNT>
F:16-99/Product: F2 chain #status predicted <F2C>
F:105-542/Product: F1 chain #status predicted <F1C>
F:105-135/Region: hydrophobic
F:486-512/Domain: transmembrane #status predicted <TMN>
F:21-53,159,397/Blinding site: carboxylate (Asn) (covalent) #status predicted

Query Match	40.6%	Score 65;	DB 2;	Length 542;
Best Local Similarity	86.7%	Pred. No. 0.31;		
Matches 13; Conservative	1;	Mismatches	1;	Indels 0;
Gaps	0;			

QY	3	LSEIKGVIVHRLGV	17
		:	
		:	
Db	280	LSEVKGIVHRLGV	294

```

RESULT 12
VGNZCD
cell fusion glycoprotein precursor - canine distemper virus
N:Contains: fusion protein F1; fusion protein F2
C:Species: canine distemper virus
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: J50321
R:Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.
Virus Res. 8, 373-386, 1987
A:Title: The nucleotide sequence of the gene encoding the F protein of canine distemper
A:Reference number: J50321; MUID:88129050
A:Accession: J50321
A:Molecule type: mRNA
A:Residues: 1-662 <BAR>
A:Cross-references: GB:M21849; NID:g323241; PIDN:AAA42878.1; PID:g323242
C:Genetics:

```

```
C:Superfamily parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-135/Domain: signal sequence #status predicted <SIG>
F:136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>
F:225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>
F:606-629/Domain: transmembrane #status predicted <MEM>
F:62.141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          40.6%; Score 65; DB 1; Length 662;
Best Local Similarity 86.7%; Pred. No. 0.33;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

RESULT 13
S21382
cell fusion protein - canine distemper virus
C:Species: canine distemper virus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999

C:Accession: S21382
R:Wild, T.F.; Bernard, A.; Spehner, D.; Villevall, D.; Drilling, R.
submitted to the EMBL Data Library, April 1992
A:Description: Vaccination of mice against canine distemper virus induced encephalitis
A:Reference number: S21382
A:Accession: S21382
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-662 <Wild>
A:Cross-references: EMBL:X65509: NID:958853: PIDN:CAA6481.1: PID:958854
C:Superfamily: parainfluenza virus cell fusion protein

Query Match	40.6%	Score 65;	DB 2;	Length 662;
Best Local Similarity	86.7%	Pred. No. 0.39;		
Matches 13; Conservative	1;	Mismatches	1;	Indels 0;
QY	3	LSEIKGVIVHRLGV	17	
		:		
DB	400	LSEKGVIVHRLGV	414	

```

RESULT 14
VCN2PD
cell fusion glycoprotein precursor - phocine distemper virus
N:Contains: fusion protein F1, fusion protein F2
C:Species: phocine distemper virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996
C:Accession: J01368
R:Keywords: J., Blixenkroge-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.
J. Gen. Virol. 72, 2859-2966, 1991
A:Title: The nucleotide sequence and deduced amino acid composition of the haemagglut
A:Reference number: J01368; MUID:92113538
A:Accession: J01368
A:Molecule type: genomic RNA
A:Residues: 1-631 <KOV>
C:Genetics:

```

C:	Superfamily: paramyxovirus	cell fusion protein
C:	Keywords: glycoprotein; membrane fusion; transmembrane protein	
F:1-188/Product:	cell fusion glycoprotein F2	#status predicted <FP2>
F:89-106/Domain:	transmembrane	#status predicted <TM1>
F:189-193/Region:	cleavage processing	#status predicted <CP>
F:194-631/Product:	cell fusion glycoprotein F1	#status predicted <FP1>
F:194-212/Domain:	transmembrane	#status predicted <TM2>
F:575-595/Domain:	transmembrane	#status predicted <TM3>
F:110_142_148_486/Binding site:	carbohydrate (Asn) (covalent)	#status predicted

Query Match	40.0%;	Score 64;	DB 1;	Length 631;
Best Local Similarity	80.0%;	Pred. No. 0.51;		
Matches 12;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;

QY	3	LSEIKGVIVHRLEGV	17
		: I	
Db	369	LSEVKGVVHRLNAV	383

RESULT 15
A48346
cell fusion glycoprotein precursor - phocine distemper virus (strain Uster/88)
N:Contains: fusion protein P1: fusion protein P2
C:Species: phocine distemper virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 25-Oct-1996
C:Accession: A48346
R:Curran, M.D.; Lu, Y.-J.; Rima, B.K.
Arch. Virol. 126, 159-169, 1992
A:Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced a
A:Reference number: A48346; MUID:92398437

A:Molecule type: mRNA
A:Residues: 1-631 <CUR>
A>Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBI:P.113099)
C:Genetics:
A:Gene: F

C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F:111-553/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F:501-517/Domain: transmembrane #status predicted <TMN>
F:32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.0%; Score 72; DB 1; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLGCV 17
Db 291 LSEIKGVIHRLGCV 305
|||||

RESULT 7
VGNRRK
cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)
N:Contains: fusion glycoprotein F1; fusion glycoprotein F2
C:Species: rinderpest virus
C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 25-Oct-1996
C:Accession: A31051
R:Hu, D.; Yamataka, M.; Miller, J.; Dale, B.; Grubman, M.; Ylma, T.
Virology 166, 149-153, 1988
A:Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis
A:Reference number: A31051; MUID:88322864
A:Accession: A31051
A:Molecule type: genomic RNA
A:Residues: 1-546 <RSD>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-108/Product: cell fusion glycoprotein F2 #status predicted <FF1>
F:109-546/Product: cell fusion glycoprotein F1 #status predicted <FF2>
F:109-134/Domain: transmembrane #status predicted <TMN>
F:481-513/Domain: transmembrane #status predicted <TN2>
F:25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.4%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.045;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLGCV 17
Db 284 LSEIKGVIHRLGCV 298
|||||

RESULT 8
S47305
gene F protein - rinderpest virus
C:Species: rinderpest virus
C>Date: 20-Oct-1994 #sequence_revision 08-Sep-1995 #text_change 20-Sep-1999
C:Accession: S47305; S47301
R:Barron, M.D.; Barrett, T.
Submitted to the EMBL Data Library, March 1994
A:Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30
A:Reference number: S47283
A:Accession: S47305
A:Molecule type: mRNA
A:Residues: 1-546 <BAR>
A:Cross-references: EMBL:Z30697; NID:G535396; PIDN:CAA83181.1; PID:G535401; EMBL:Z30700;
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: transmembrane protein

Query Match 44.4%; Score 71; DB 2; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.045;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLGCV 17
|||||

Db 284 LSEIKGVIHRLGCV 298

RESULT 9
S47034
cell fusion protein precursor - porpoise morbillivirus
N:Alternate names: F protein
C:Species: porpoise morbillivirus
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S47034
R:Bolt, G.; Gottschalk, E.; Blixenkron-Moeller, M.; Wilsaupt, R.G.A.; Welsh, M.J.;
submitted to the EMBL Data Library, July 1994
A:Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbi
A:Reference number: S47034
A:Accession: S47034
A:Molecule type: mRNA
A:Residues: 1-552 <BOL>
A:Cross-references: EMBL:X80757; NID:G520639; PIDN:CAA56731.1; PID:G520640
A:Experimental source: isolated Uster 88
A>Note: The source is designated as Cetacean morbillivirus
C:Superfamily: parainfluenza virus cell fusion protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-552/Product: fusion protein #status predicted <MAT>

Query Match 41.6%; Score 66.5; DB 2; Length 552;
Best Local Similarity 61.5%; Pred. No. 0.2;
Matches 16; Conservative 1; Mismatches 2; Indels 7; Gaps 1;

OY 3 LSEIKGVIHRLGCVPSLHMSYGL 28
Db 290 LSEIKGVIHRLGCVPSLHMSYGL 308
|||||

RESULT 10
VGNRL
cell fusion glycoprotein precursor - rinderpest virus (strain L)
N:Contains: fusion glycoprotein F1; fusion glycoprotein F2
C:Species: rinderpest virus
C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C:Accession: A28921
R:Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.
Virology 164, 523-530, 1988
A:Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of t
A:Reference number: A28921; MUID:88219541
A:Accession: A28921
A:Molecule type: mRNA
A:Residues: 1-546 <TSU>
A:Cross-references: GB:M20870; NID:G333898; PIDN:AAA47399.1; PID:G333899
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-104/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F:105-546/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F:109-133/Domain: transmembrane #status predicted <TMN>
F:485-513/Domain: transmembrane #status predicted <TN2>
F:25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.2%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.23;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 LSEIKGVIHRLGCV 17
Db 284 LSEIKGVIHRLGCV 298
|||||

RESULT 11
JQ2223
cell fusion protein F0 precursor - phocine distemper virus
N:Contains: F1 and F2 chains
C:Species: phocine distemper virus


```
RESULT 3
JU0274
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain X
N:Contains: fusion glycoprotein F1; fusion glycoprotein F2
C:Species: subacute sclerosing panencephalitis virus, SSPV
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
C:Accession: JU0274
R:Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.
Virus Genes 4, 173-181, 1990
A>Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe
A:Reference number: JU0274; MUID:90385702
A:Accession: JU0274
A:Molecule type: mRNA
A:Residues: 1-534 <KOM>
A:Cross-references: EMBL:DJ0548; NID:g222256; PIDN:BA01405.1; PID:g222257
A>Note: the authors translated the codon GTA for residue 439 as Gly and GGG for residue
C:Genetics:
A:Gene: F
C:Superfamily: paraInfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F:108-534/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F:498-514/Domain: transmembrane #status predicted <TMN>
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          45.0%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVVHRLEGV 17
|||||
Db 288 LSEIKGVVHRLEGV 302

RESULT 4
S47300
gene F protein - rinderpest virus
C:Species: rinderpest virus
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: S47300; PQ0865
R:Evans, S.A.; Barton, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.
submitted to the EMBL data library, March 1994
A:Description: The complete nucleotide sequence of the fusion protein gene of the vacct
A:Reference number: S47299
A:Accession: S47300
A:Molecule type: DNA
A:Residues: 1-546 <EVA>
A:Cross-references: EMBL:Z31656; NID:g535406; PIDN:CA083482.1; PID:g535407
R:Chamberlain, R.W.; Manwayi, H.M.; Hockley, E.; Shalla, M.S.; Goatley, L.; Knowles, N.J
J. Gen. Virol. 74, 2775-2780, 1993
A>Title: Evidence for different lineages of rinderpest virus reflecting their geographic
A:Reference number: PQ0865; MUID:94103786
A:Accession: PQ0865
A:Molecule type: mRNA
A:Residues: 86-191 <CHA>
C:Genetics:
A:Gene: F
C:Superfamily: paraInfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
C:Accession: PQ0865; MUID:94103786

Query Match          45.0%; Score 72; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVVHRLEGV 17
|||||
Db 284 LSEIKGVVHRLEGV 298

RESULT 5
```

```
E48556
cell fusion glycoprotein precursor - measles virus (strain AIR-C)
C:Species: measles virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
C:Accession: E48556
R:Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.
Virus Genes 7, 67-81, 1993
A>Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIR
A:Reference number: A48556; MUID:93227570
A:Accession: E48556
A:Molecule type: genomic RNA
A:Residues: 1-550 <MOR>
A:Cross-references: GB:S58435; NID:g299460; PIDN:AA026145.1; PID:g299465
A>Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBIPI:129272)
C:Genetics:
A:Gene: F
C:Superfamily: paraInfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F:108-550/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F:113-138/Region: hydrophobic
F:495-514/Domain: transmembrane #status predicted <TMN>
F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          45.0%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LSEIKGVVHRLEGV 17
|||||
Db 288 LSEIKGVVHRLEGV 302

RESULT 6
VGNZMY
cell fusion glycoprotein precursor - measles virus
C:Species: measles virus
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000
C:Accession: A26962; A2616; PQ0384
R:Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.
J. Gen. Virol. 68, 1695-1703, 1987
A>Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and co
A:Reference number: A92794; MUID:87224816
A:Accession: A26962
A:Molecule type: mRNA
A:Residues: 1-553 <BNC>
A:Cross-references: GB:D00090; NID:g222061; PIDN:BA00056.1; PID:g222062
A:Experimental source: Strain Halle
R:Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Bellint
Virology 155, 508-523, 1986
A>Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles v
A:Reference number: A94350; MUID:87071668
A:Accession: A2616
A:Molecule type: mRNA
A:Residues: 4-553 <RIC>
A:Cross-references: GB:M14915; NID:g331762; PIDN:AAA46423.1; PID:g331763
A:Experimental source: strain Edmonston
R:Schultz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A>Title: A measles virus isolate from a child with Kawasaki disease: sequence compari
A:Reference number: PQ0374; MUID:92300360
A:Accession: PQ0380
A:Molecule type: genomic RNA
A:Residues: 272-553 <SCH>
A:Experimental source: Isolate CL
A:Accession: PQ0384
A:Molecule type: genomic RNA
A:Residues: 272-553 <SCH2>
A:Experimental source: Isolate SE
C:Genetics:
A:Gene: F
C:Superfamily: paraInfluenza virus cell fusion protein
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 7.22009 Seconds

(Without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-9

Perfect score: 160

Sequence: 1 KLSEIKGVIVHRLGVGSPSLHMSYGLRFX 31

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	45.0	282	2 P00376	cell fusion glycop
2	72	45.0	282	2 P00388	cell fusion glycop
3	72	45.0	534	1 J00274	cell fusion glycop
4	72	45.0	546	2 S47300	gene F protein - r
5	72	45.0	550	1 E48556	cell fusion glycop
6	72	45.0	553	1 VGNZMV	cell fusion glycop
7	71	44.4	546	1 VGNZRK	cell fusion glycop
8	71	44.4	546	2 S47305	gene F protein - r
9	66.5	41.6	552	2 S47034	cell fusion protei
10	66	41.2	546	1 VGNZRL	cell fusion glycop
11	65	40.6	542	2 J02223	cell fusion protei
12	65	40.6	662	1 VGNZCD	cell fusion glycop
13	65	40.6	662	2 S21382	cell fusion protei
14	64	40.0	631	1 VGNZPD	cell fusion glycop
15	64	40.0	631	1 A48346	cell fusion glycop
16	60	37.5	546	2 S55386	cell fusion protei
17	57.5	35.9	92	1 RHHUG	gonadoliberin prec
18	57	35.6	90	1 RHMSG	gonadoliberin prec
19	57	35.6	92	1 RHRTG	gonadoliberin prec
20	56	35.0	636	2 S47299	gene F protein - r
21	54	33.8	67	2 I78541	gonadoliberin prec
22	53	33.1	508	1 VGVNFR	splike glycoprotein
23	52	32.5	10	1 RHPGG	gonadoliberin - pi
24	52	32.5	10	1 RHSHG	gonadoliberin - sh
25	52	32.5	89	2 I51423	gonadoliberin prec
26	52	32.5	379	2 DCBSPK	phosphoribosylam
27	50.5	31.6	190	2 T37168	probable tetra-fam
28	50	31.2	333	2 T23151	hypothetical prote
29	49	30.6	451	2 AH0063	conserved hypotnet

30	49	30.6	582	2 G71500	aspartate--trna 11
31	48.5	30.3	393	2 D83589	glutaryl-coa dehyd
32	48.5	30.3	413	2 B70907	hypothetical prote
33	48	30.0	10	1 RHM01	gonadoliberin I -
34	48	30.0	92	2 I50644	gonadoliberin I pr
35	48	30.0	98	2 I50739	gonadoliberin I pr
36	48	30.0	502	2 I36589	probable transmemb
37	48	30.0	576	2 AF2361	flavoprotein (limpo
38	48	30.0	856	2 I58411	protein-tyrosine k
39	47.5	29.7	222	1 E71024	probable transport
40	47.5	29.7	229	2 JG7219	nuclear protein SR
41	47.5	29.7	586	2 T08293	hypothetical prote
42	47	29.4	339	2 B97755	hypothetical prote
43	47	29.4	388	2 C72710	probable rnu prote
44	47	29.4	452	2 T10882	coproporphyrinogen
45	47	29.4	516	2 T00514	cytochrome P450 ho

ALIGNMENTS

RESULT 1

P00376 cell fusion glycoprotein - measles virus (strain TT) (fragment)

C:Species: measles virus

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999

C:Accession: P00376

R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compari

A:Reference number: P00374; MUID:92300360

A:Accession: P00376

A:Molecule type: genomic RNA

A:Residues: 1-282 <SCH>

C:Genetics:

A:Superfamily: paramyxo-influenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion

Query Match 45.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17

DB 20 LSEIKGVIVHRLGV 34

RESULT 2

P00388 cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)

C:Species: measles virus

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999

C:Accession: P00388

R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.

J. Gen. Virol. 73, 1581-1586, 1992

A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compari

A:Reference number: P00374; MUID:92300360

A:Accession: P00388

A:Molecule type: genomic RNA

A:Residues: 1-282 <SCH>

C:Genetics:

A:Superfamily: paramyxo-influenza virus cell fusion protein

C:Keywords: glycoprotein; membrane fusion

Query Match 45.0%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17

DB 20 LSEIKGVIVHRLGV 34


```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 64098/102/ARDE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-000-931-5

Query Match          96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 HWSYGLRP 9
Db      2 HWSYGLRP 9

RESULT 15
US-08-428-488-22
Sequence 22, Application US/08428488
Patent No. 5624894
GENERAL INFORMATION:
APPLICANT: BODOR, Nicholas S.
TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,488
FILING DATE: 27-APR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Mary Katherine
REGISTRATION NUMBER: 26,254
REFERENCE/DOCKET NUMBER: 028724-087
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-6620
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Position 1 = p-Glu."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10

```

```

OTHER INFORMATION: /note= "Position 10 = Gly-NH2."
US-08-428-488-22

Query Match          96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 HWSYGLRP 9
Db      2 HWSYGLRP 9

Search completed: October 10, 2002, 16:14:00
Job time : 3.11538 secs

```

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,935
FILING DATE: 24-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 547670man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 846-286-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note= "C-terminal amide"
US-08-184-935-6

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 13
US-08-343-883-1
Sequence 1, Application US/08343883
Patent No. 5573767
GENERAL INFORMATION:
APPLICANT: Dufour, Raymond J.
APPLICANT: Roulet, Claude J.M.
APPLICANT: Chouvet, Claire D.
APPLICANT: Bonneau, Michel B.
TITLE OF INVENTION: Method for improving the organoleptic
TITLE OF INVENTION: qualities of the meat from uncastrated male domestic
TITLE OF INVENTION: animals, vaccines which are usable in this method, new
TITLE OF INVENTION: peptide, in particular for producing these vaccines...
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Larson and Taylor
STREET: 727 Twenty-third Street, South
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,883
FILING DATE: 17-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/946,495
FILING DATE: 09-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9102513
FILING DATE: 01-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9115289
FILING DATE: 10-DEC-1991
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 10
OTHER INFORMATION: /label= NH2
OTHER INFORMATION: /note= "amidated glycine"
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: /label= pyro
OTHER INFORMATION: /note= "pyroglutamic acid"
PUBLICATION INFORMATION:
AUTHORS: Matsuo, H.
AUTHORS: Baba, Y.
AUTHORS: G. Nair, R. M.
AUTHORS: Arimura, A.
AUTHORS: Schally, A. V.
TITLE: Structure of the porcine LH- and
TITLE: FSH-releasing hormone. I. The proposed amino acid
TITLE: sequence.
JOURNAL: Biochem. Biophys. Res. Commun.
VOLUME: 43
ISSUE: 6
PAGES: 1334-1339
DATE: 1971
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 10
US-08-343-883-1

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 14
US-08-000-931-5
Sequence 5, Application US/08000931
Patent No. 5578477
GENERAL INFORMATION:
APPLICANT: Tamanoi Dr., Fuyuhiko
TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF
TITLE OF INVENTION: INHIBITORS OF PROTEIN FARNESYLTRANSFERASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/000,931
FILING DATE: 05-JAN-1994

TOPOLGY: linear
MOLECULE TYPE: protein
US-07-690-983D-37

Query Match
Best Local Similarity 96.3%; Score 52; DB 1; Length 10;
100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
|||||||
DB 3 HWSYGLRP 10

RESULT 10
US-08-103-022-1
Sequence 1, Application US/08103022
Patent No. 5413990
GENERAL INFORMATION:
APPLICANT: Haviv, Fortuna
APPLICANT: Fitzpatrick, Timothy D.
APPLICANT: Swenson, Kolf E.
APPLICANT: Nichols, Charles J.
APPLICANT: Mort, Nicholas A.
TITLE OF INVENTION: N-Terminus Modified Analogs of LHRH
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377
STREET: Abbott Laboratories, One Abbott Park Road
CITY: No. 5413990th Chicago
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,022
FILING DATE: 05-OCT-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 5389.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 938-2623
OTHER INFORMATION: /note="Xaa at position 1 is a
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="Xaa at position 1 is a
OTHER INFORMATION: 5-oxo-prolyl aminoacyl residue."
US-08-103-022-1

Query Match
Best Local Similarity 96.3%; Score 52; DB 1; Length 10;
100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
|||||||
DB 2 HWSYGLRP 9

RESULT 11
US-07-897-680-1

Sequence 1, Application US/07897680
Patent No. 5446025
GENERAL INFORMATION:
APPLICANT: Fu Lu, Mou-Ying
APPLICANT: Subba Rao, Gowdahalain N.
APPLICANT: Lee, Dennis Y.
TITLE OF INVENTION: Formulations and Method for the
TITLE OF INVENTION: Percutaneous Administration of Leuprolide
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dept. 377 Abbott Laboratories
STREET: One Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/897,680
FILING DATE: 19920612
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 5165.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 937-9556
OTHER INFORMATION: /note="XAA at position 1 is a
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="XAA at position 1 is a
OTHER INFORMATION: pyro-glutamyl residue"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note="XAA at position 10 is a
OTHER INFORMATION: glycyl-amide residue"
US-07-897-680-1

Query Match
Best Local Similarity 96.3%; Score 52; DB 1; Length 10;
100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
|||||||
DB 2 HWSYGLRP 9

RESULT 12
US-08-184-935-6
Sequence 6, Application US/08184935
Patent No. 5476770
GENERAL INFORMATION:
APPLICANT: PRADLES, PHILIPPE
TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN
TITLE OF INVENTION: OR HAPTEN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: P. C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400

US-07-690-983D-8
; Sequence 8, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "Represents glycinamide"
US-07-690-983D-8
Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9
RESULT 8
US-07-690-983D-32
; Sequence 32, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-690-983D-32
Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9
RESULT 9
US-07-690-983D-37
; Sequence 37, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690, 983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-07-690-983D-2
```

```
Query Match          96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 HWSYGLRP 9
        |||||
Db      2 HWSYGLRP 9
```

```
RESULT 5
US-07-690-983D-6
Sequence 6, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690, 983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
```

```
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note="Represents glycylamide"
US-07-690-983D-6
```

```
Query Match          96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 HWSYGLRP 9
        |||||
Db      2 HWSYGLRP 9
```

```
RESULT 6
US-07-690-983D-7
Sequence 7, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690, 983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note="Represents glycylamide"
US-07-690-983D-7
```

```
Query Match          96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      2 HWSYGLRP 9
        |||||
Db      2 HWSYGLRP 9
```

```
RESULT 7
```


RESULT 2
US-07-983-111-1
; Sequence 1, Application US/07983111
; Patent No. 5284657
; GENERAL INFORMATION:
; APPLICANT: Fu Lu, Mou-Ying
; APPLICANT: Reiland, Thomas L.
; TITLE OF INVENTION: "Compositions and Methods for the
; TITLE OF INVENTION: Sublingual or Buccal Administration of Therapeutic Agents"
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman, Jr., Dept. 377-AP6D
; STREET: One Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/983.111
; FILING DATE: 30-NOV-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/750.843
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Janssen, Jerry F.
; REGISTRATION NUMBER: 29,175
; REFERENCE/DOCKET NUMBER: 4848, US. 01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 938-7742
; TELEFAX: (708) 937-9556
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa at position 1 is
; OTHER INFORMATION: 5-oxo-proline"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "Xaa at position 10 is
; OTHER INFORMATION: glycynamide"
US-07-983-111-1
Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9
RESULT 3
US-07-690-983D-1
; Sequence 1, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.

APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690.983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5399
; TELEFAX: (202)672-5300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Represents pyroglutamic
; OTHER INFORMATION: acid"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "Represents glycynamide"
US-07-690-983D-1
Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9
RESULT 4
US-07-690-983D-2
; Sequence 2, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 2.11538 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-1

Perfect score: 54

Sequence: 1 XHMSYGLRPX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 segs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents:AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	96.3	10	1	US-07-714-540-9
2	52	96.3	10	1	US-07-983-111-1
3	52	96.3	10	1	US-07-690-983D-1
4	52	96.3	10	1	US-07-690-983D-2
5	52	96.3	10	1	US-07-690-983D-6
6	52	96.3	10	1	US-07-690-983D-7
7	52	96.3	10	1	US-07-690-983D-8
8	52	96.3	10	1	US-07-690-983D-32
9	52	96.3	10	1	US-07-690-983D-37
10	52	96.3	10	1	US-08-103-022-1
11	52	96.3	10	1	US-07-897-680-1
12	52	96.3	10	1	US-08-184-935-6
13	52	96.3	10	1	US-08-343-883-1
14	52	96.3	10	1	US-08-000-931-5
15	52	96.3	10	1	US-08-428-488-22
16	52	96.3	10	1	US-08-341-219-11
17	52	96.3	10	1	US-08-453-588-2
18	52	96.3	10	1	US-08-453-588-4
19	52	96.3	10	1	US-08-453-588-6
20	52	96.3	10	1	US-08-453-588-8
21	52	96.3	10	1	US-08-453-588-10
22	52	96.3	10	1	US-08-453-588-12
23	52	96.3	10	1	US-08-453-588-14
24	52	96.3	10	1	US-08-453-588-16
25	52	96.3	10	1	US-08-453-588-19
26	52	96.3	10	1	US-08-453-588-22
27	52	96.3	10	1	US-08-188-223-1

28	52	96.3	10	1	US-08-188-223-3	Sequence 3, Appl
29	52	96.3	10	1	US-08-188-223-8	Sequence 8, Appl
30	52	96.3	10	1	US-08-406-935-5	Sequence 5, Appl
31	52	96.3	10	1	US-08-591-917-1	Sequence 1, Appl
32	52	96.3	10	1	US-08-387-156-2	Sequence 2, Appl
33	52	96.3	10	1	US-08-474-555-1	Sequence 1, Appl
34	52	96.3	10	1	US-08-446-692-1	Sequence 1, Appl
35	52	96.3	10	1	US-08-242-678D-1	Sequence 1, Appl
36	52	96.3	10	2	US-08-796-598-6	Sequence 6, Appl
37	52	96.3	10	2	US-08-694-865-2	Sequence 2, Appl
38	52	96.3	10	2	US-08-694-865-18	Sequence 18, Appl
39	52	96.3	10	2	US-08-488-351A-1	Sequence 1, Appl
40	52	96.3	10	2	US-08-480-494B-1	Sequence 1, Appl
41	52	96.3	10	2	US-08-447-175A-6	Sequence 6, Appl
42	52	96.3	10	2	US-08-747-137-116	Sequence 116, App
43	52	96.3	10	2	US-08-878-748-2	Sequence 2, Appl
44	52	96.3	10	3	US-08-521-079-2	Sequence 2, Appl
45	52	96.3	10	3	US-08-521-079-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-07-714-540-9
Sequence 9, Application US/0714540
Patent No. 5262521
GENERAL INFORMATION:
APPLICANT: Almqvist, Ronald G.
TITLE OF INVENTION: ISOLATED ATRIAL PEPTIDE-DEGRADING
TITLE OF INVENTION: ENZYME AND NOVEL COMPOUNDS USEFUL AS INHIBITORS THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Irell & Manella
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/714,540
FILING DATE: 19910607
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Dianne E.
REGISTRATION NUMBER: 31,292
REFERENCE/DOCKET NUMBER: 8500-0135.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-714-540-9

Query Match 96.3% Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HMSYGLRP 9
| | | | | | | |
Db 2 HMSYGLRP 9

XX Synthetic.
 XX EP143573-A.
 XX
 XX PD 05-JUN-1985.
 XX
 XX PF 05-NOV-1984; 84EP-0307625.
 XX
 XX PR 29-NOV-1983; 83US-0556148.
 XX PR 30-AUG-1985; 85US-0771517.
 XX
 XX PA (SALK) SALK INST FOR BIOL. STUD.
 XX
 XX PI Roeske RW, Rivier JE, Vale WM;
 XX
 XX DR WPI; 1985-136434/23.
 XX
 XX PT New GnRH antagonist peptide(s) - useful as inhibitors of
 XX gonadotropin(s) and/or steroid(s) for contraceptive use.
 XX
 XX PS Disclosure; Page 1; 20pp; English.
 XX
 XX CC The claimed peptide antagonists inhibit the release of gonadotrophins
 XX and/or steroids. They are antagonistic to GnRH, inhibit ovulation, and
 XX may cause resorption of a fertilised egg if administered shortly after
 XX absorption. The peptides also have utility in male contraception, and
 XX in treatment of precocious puberty, hormone dependent neoplasia,
 XX CC dysmenorrhoea and endometriosis.
 XX
 XX SQ Sequence 10 AA:

Query Match 96.3%; Score 52; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||||
 DB 2 HWSYGLRP 9

RESULT 14
 AAP60127
 ID AAP60127 standard; Peptide; 10 AA.
 XX
 XX AC AAP60127;
 XX
 XX DT 12-JUN-1991 (first entry)
 XX
 XX DE Gonadoliberin antagonist.
 XX
 XX KW Gonadoliberin antagonist; contraceptive; antitumor.
 XX
 XX PN EP201260-A.
 XX
 XX PD 12-NOV-1986.
 XX
 XX PF 28-APR-1986; 86EP-0303210.
 XX
 XX PR 09-MAY-1985; 85US-0732531.
 XX
 XX PA (SALK) SALK INST FOR BIOL. STUD.
 XX
 XX PI Rivier JEF, Varga JI, Hagler AT, Struthers RS, Perrin MH;
 XX PI Rivier CL, Vale WM;
 XX
 XX DR WPI; 1986-299774/46.
 XX
 XX PT New peptide gonadotropin releasing hormone antagonists - useful
 XX esp. as contraceptives, for treating early puberty,
 XX hormone-dependent neoplasms etc.
 XX
 XX PS Disclosure; Page 1; 33pp; English.

XX
 CC The decapeptide encodes a gonadoliberin antagonist, which may be
 CC used as a male contraceptive and as an antitumor (against steroid-
 CC dependent tumours).
 XX

SQ Sequence 10 AA:

Query Match 96.3%; Score 52; DB 7; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||||
 DB 2 HWSYGLRP 9

RESULT 15
 AAP61403
 ID AAP61403 standard; protein; 10 AA.
 XX
 XX AC AAP61403;
 XX
 XX DT 04-AUG-1991 (first entry)
 XX
 XX DE Gonadotropin releasing hormone.
 XX
 XX KW Gonadotropin releasing hormone; analogue; peptide synthesis;
 XX ovulation; veterinary medicine; fertility;
 XX
 XX PN DD232500-A.
 XX
 XX PD 29-JAN-1986.
 XX
 XX PF 08-MAY-1984; 84DD-0262804.
 XX
 XX PR 08-MAY-1984; 84DD-0262804.
 XX
 XX PA (DEAK) AKAD WISSENSCHAFT DDR.
 XX
 XX PI Kaufmann KD, Dolling R, Handel L;
 XX
 XX DR WPI; 1986-137868/22.
 XX
 XX PT Prepn. of gonadotropin liberating hormone and analogues - by
 XX PT multistage rapid peptide synthesis in soln. without isolating
 XX PT intermediates
 XX
 XX PS Disclosure; page 7; 8pp; german.
 XX
 XX CC The gonadotropin releasing hormone and its analogues are prepd. by a
 XX CC new multistage rapid peptide synthesis method in soln., where the
 XX CC intermediates are not isolated. The process is rapid and gives very
 XX CC pure peptide quickly and using little equipment. The peptide can be
 XX CC used in veterinary medicine to synchronise ovulation in large animal
 XX CC herds, and in human medicine in the treatment of fertility disorders.
 XX

SQ Sequence 10 AA:

Query Match 96.3%; Score 52; DB 7; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||||
 DB 2 HWSYGLRP 9

Search completed: October 10, 2002, 16:05:06
 Job time : 7.25641 secs

XX Sequence 10 AA;
 SQ Query Match 96.3%; Score 52; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||
 Db 2 HWSYGLRP 9

RESULT 11
 AAP10416
 ID AAP10416 standard; peptide; 10 AA.
 XX AAP10416;
 AC
 XX 17-DEC-1992 (first entry)
 DT
 XX
 DE Lutetinsing Hormone Releasing Hormone analogue #5.
 XX
 KW LHRH; Follicle Stimulating Factor; FSH; acne; hirsutism;
 KW dysmenorrhea; precocious puberty; endometriosis; prostate cancer;
 KW benign prostate hypertrophy; mammary tumour.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /label= OTHER
 FT /note= "pyroglutamic acid"
 FT Modified-site 7 /label= OTHER
 FT /note= "N-alpha-methyl-Leu"
 FT Modified-site 10 /note= "amidated or absent, in which case Pro(9)
 FT is Pro-NH-C2H5"
 FT
 FT
 XX BE885308-A.
 PN
 XX 19-MAR-1981.
 PD
 XX 23-FEB-1983; 83BE-0468932.
 PF
 XX 21-SEP-1979; 79FR-0023545.
 PR
 XX (ROUS) ROUSSEL UCLAF.
 PA
 XX WPI: 1981-23409D/14 (23409D).
 DR
 XX LH-RH, liberating factor for LH and FSH, and its agonists compsn.
 PT - used to treat prostate adenocarcinoma, benign hypertrophy of
 PT the prostate, hirsutism, acne, etc.
 PT
 XX
 PS Claim 1(f); Page 16; 27pp; French.
 XX
 CC A composition is claimed containing LHRH or its analogues. The
 CC composition is used to treat prostate adenocarcinoma, benign
 CC hypertrophy of the prostate, endometriosis, dysmenorrhea, hirsutism,
 CC hormone-dependent mammary tumours, for treatment or prevention of
 CC precocious puberty, delaying the onset of puberty and for treating
 CC acne. The compositions may also contain antiandrogens.
 CC See AAP10411-P10418.
 CC
 SQ Sequence 10 AA;

Query Match 96.3%; Score 52; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||
 Db 2 HWSYGLRP 9

v

RESULT 12
 AAP20277
 ID AAP20277 standard; Protein; 10 AA.
 XX AAP20277;
 AC
 XX 30-NOV-1992 (first entry)
 DT
 XX
 DE Modified carboxy terminal peptide 2.
 DE
 XX
 KW Medicament; pituitary function; hypothalamic releasing factors;
 KW enkephalin.
 KW
 XX
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 10 /label= Modified_Gly_with_terminal_CHN2_or_CH2X
 FT /note= "X= Cl, Br, I"
 FT
 FT
 XX US4305872-A.
 PN
 XX 15-DEC-1981.
 PD
 XX 19-OCT-1979; 79US-0086417.
 PF
 XX 19-OCT-1979; 79US-0086417.
 PR
 XX 19-OCT-1979; 79US-0086417.
 PA (WING/) WINGROVE K.
 XX
 PI Johnston RB, Balk JT, Pelton JT;
 PI
 DR WPI: 1982-01722E/01 (01722E).
 XX
 XX Diazo- and halo-methyl ketone derivs. of polypeptide(s) - useful
 PT as hormone or opiate antagonists or agonists
 PT
 XX
 PS Claim 2; Page 25; 26pp; English.
 XX
 CC The sequences given in AAP20276-80 are biologically active derivatives
 CC of carboxy-terminal polypeptides which have the OH group of the
 CC terminal carboxy group replaced by CHN2 or CH2X (where X = Cl, Br or
 CC I). These peptides are agonists or antagonists to the polypeptides
 CC from which they are derived. They can be used as medicaments eg.
 CC for regulating pituitary function, or esp. as research tools for
 CC investigating the action of the polypeptides in biological systems.
 CC The polypeptides used to obtain the derivatives were selected from
 CC acid free forms of hypothalamic releasing factors, enkephalins and
 CC biologically active polypeptide fragments of these.
 CC
 SQ Sequence 10 AA;

Query Match 96.3%; Score 52; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||
 Db 2 HWSYGLRP 9

RESULT 13
 AAP50222
 ID AAP50222 standard; Protein; 10 AA.
 XX AAP50222;
 AC
 XX 20-JAN-1992 (first entry)
 DT
 XX
 DE Gonadotrophin release stimulating hormone.
 DE
 XX
 KW GnRH; LH-RH; LRF; gonadotrophins; steroids; contraceptive.

PR	03-JUN-1999;	99GB--0012807.
PR	03-JUN-1999;	99US-0137592.
XX		
PA	(BIOI-) BIO INNOVATION LTD.	
P1	Franks CR, Delia Bitta R, Maitland NJ, Knight DJ;	
DR	WPI: 2001-061658/07.	
XX		
PT	Novel product comprising proliferatively active moiety linked to	
PT	genetic material, useful as vectors for protected nucleic acid material	
PT	and as mitogen to stimulate proliferation of target cell -	
XX		
PS	Disclosure: Page 4; 49pp; English.	
CC		
CC	The present invention relates to a product comprising a proliferatively	
CC	active moiety (PAM) linked to nucleic acid material which is associated	
CC	with a protective material. The PAM product is useful for manufacturing	
CC	a medicament for treating e.g. an autoimmune disease, transplant	
CC	rejection, retroviral disease, graft-versus-host-disease, or	
CC	lymphoproliferative disease, comprising cells bearing a high affinity	
CC	receptor for PAM. The present sequence is a peptide of	
CC	gonadotropin-releasing hormone (GnRH). GnRH is a peptide hormone, which	
CC	has high-affinity receptors, and therefore can be used in the present	
CC	invention.	
SQ	Sequence 9 AA:	
Query Match :	96.3%; Score 52; DB 22; Length 9;	
Best Local Similarity	100.0%; Pred. No. 6.4e+05;	
Matches 8; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	2 HWSYGLRP 9	
Db	1 HWSYGLRP 8	
RESULT 9		
AAP10097		
ID	AAP10097 standard; peptide: 10 AA.	
XX		
AC	AAP10097;	
XX		
D7	19-AUG-1992 (first entry)	
XX		
DE	Sequence of luteinising hormone (LH-RH, ICSS) liberating hormone.	
XX		
KW	Gonadorelin; luteinising hormone releasing hormone; LH-RH;	
RW	ICSH; prostatic hyperplasia therapy.	
XX		
OS	Mammal.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 1	/label= Pyr
FT	Modified-site 10	
FT		/label=Gly-NH2
XX		
PN	BE887639-A.	
XX		
PD	24-AUG-1981.	
XX		
PE	27-AUG-1981; 81BE-0303944.	
XX		
PR	22-MAY-1980; 80US-0152241.	
XX		
PA	(AMHP) AYERST MCKENNA HARR.	
XX		
PI	Aucclair C;	
XX		
DR	WPI: 1981-66067D/37 (66067D).	
XX		
PT	Gonadorelin for treatment of benign prostatic hyperplasia - is	

PT	the deca:peptide Pyr-His-Trip-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH2 or
PT	Luteinising hormone liberating hormone
PS	Claim 1; Page 7; 9pp; French.
XX	
XX	The inventors claim a compsn. for the redn. or prevention of
CC	undesired prostatic growth in males. The compsn. contains a
CC	decapeptide (gonadorelin) (AAP10097) with an appropriate vehicle or
CC	support. The compsn. is used for treating e.g. benign prostatic
CC	hyperplasia by parenteral admin. in daily doses of 0.035-11.0 (pref.
CC	0.080-2.0) mg/kg. Gonadorelin is the generic name for LH-RH and is
CC	described in US3835108. In the example s.c. injection of the
CC	compsn. significantly reduced the wt. of seminal vesicles and
CC	ventral prostate in rats without affecting the wt. of the
CC	testicles.
SQ	
XX	Sequence 10 AA;
XX	
XX	Query Match 96.3%; Score 52; DB 2; Length 10;
XX	Best Local Similarity 100.0%; Pred. No. 0.0039;
XX	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	
DB	2 HMSYGLRP 9
	2 HMSYGLRP 9
RESULT 10	
ID	AAP10411 standard; peptide; 10 AA.
AC	AAP10411;
XX	
DT	17-DEC-1992 (first entry)
XX	
DE	Luteinising Hormone Releasing Hormone.
XX	
KW	LHRH; Follicle Stimulating Factor; FSH; acne; hirsutism;
KW	dysmenorrhea; precocious puberty; endometriosis; prostate cancer;
KW	benign prostate hypertrophy; mammary tumour.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1
FT	/label= OTHER
FT	/note= "pyroglutamic acid"
FT	Modified-site 10
FT	/note= "amidated"
XX	
PN	BE885308-A.
XX	
PD	19-MAR-1981.
XX	
PP	23-FEB-1983; 83BE-0468932.
XX	
PR	21-SEP-1979; 79FR-0023545.
XX	
PA	(ROUS) ROUSSEL UCLAF.
XX	
DR	WPI: 1981-23409D/14 (23409D).
XX	
PT	LH-RH, liberating factor for LH and FSH, and its agonists compsn.
PT	- used to treat prostate adenocarcinoma, benign hypertrophy of
XX	the prostate, hirsutism, acne, etc.
PS	
XX	Claim 1(a): Page 15; 27pp; French.
XX	
CC	A composition is claimed containing LHRH or its analogues. The
CC	composition is used to treat prostate adenocarcinoma, benign
CC	hypertrophy of the prostate, endometriosis, dysmenorrhea, hirsutism,
CC	hormone-dependent mammary tumours, for treatment or prevention of
CC	precocious puberty, delaying the onset of puberty and for treating
CC	acne. The compositions may also contain antiandrogens.
CC	See also AAP10412-PI0418.

[illegible]

PF	17-MAY-2000;	2000MO-USJ3576.	
XX			
PR	17-MAY-1999;	99US-0134406.	
PR	10-SEP-1999;	99US-0153406.	
PR	15-OCT-1999;	99US-0159783.	
XX			
PA	(CONU-)	CONUCHEM INC.	
XX			
PI	Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;		
XX			
DR	WPI: 2001-112059/12.		
XX			
PT	Modifying and attaching therapeutic peptides to albumin prevents		
PT	peptidase degradation, useful for increasing length of in vivo activity		
PS	Disclosure; Page 240; 733pp; English.		
XX			
CC	The present invention describes a modified therapeutic peptide (I)		
CC	comprising a therapeutically active amino acid region (III) and a		
CC	reactive group (II) (e.g. succinimidyl and maleimido groups) attached to		
CC	a less therapeutically active amino acid region (IV), which covalently		
CC	bonds with amino/hydroxyl/thiol groups on blood components to form a		
CC	peptidase stabilised therapeutic peptide composed of 3-50 amino acids.		
CC	(I) are useful for modifying therapeutic peptides e.g. hormones, growth		
CC	factors and neurotransmitters, to protect them from peptidase activity		
CC	in vivo for the treatment of various disorders. Endogenous therapeutic		
CC	peptides are not suitable as drug candidates as they require frequent		
CC	administration due to rapid degradation by peptidases in the body.		
CC	Modifying and attaching therapeutic peptides to albumin prevents or		
CC	reduces the action of peptidases to increase length of activity (half		
CC	life) and specifically as bonding to large molecules decreases		
CC	intracellular uptake and interference with physiological processes.		
CC	AA90829 to AA92441 represent peptides which can be used in the		
CC	exemplification of the present invention.		
XX			
SO	Sequence	9 AA:	
XX			
Query Match	96.3%;	Score 52;	DB 22; Length 9;
Best Local Similarity	100.0%;	Pred. No. 6.4e+05;	
Matches	8; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
OY	2	HMSYGLRP	9
DB	1	HMSYGLRP	8
XX			
RESULT 8			
AAAB59836			
ID	AAAB59836	standard; Peptide;	9 AA.
XX			
AC	AAAB59836;		
XX			
DT	26-MAR-2001	(first entry)	
XX			
DE	GnRH peptide.		
XX			
KW	GnRH-III; autoimmune disease; transplant rejection; retroviral disease;		
KW	gnrh-versus-host-disease; lymphoproliferative disease;		
KW	gonadotropin-releasing hormone.		
XX			
OS	Petromyzon marinus.		
XX			
FT	Key	Location/Qualifiers	
FT	Modified-site	1	
FT	/note=	"Linked to Glucagon-like peptide"	
XX			
PN	WO200074724-A2.		
XX			
PD	14-DEC-2000.		
XX			
PF	05-JUN-2000;	2000MO-GB02014.	
XX			

Query Match 96.3%; Score 52; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||
 Db 1 HWSYGLRP 8

RESULT 4

AAB15363 standard; peptide: 9 AA.

AC AAB15363;
 DT 17-JAN-2001 (first entry)

DE Human LHRH peptide SEQ ID NO: 2.

KW Human, LHRH; GnRH; luteinising hormone releasing hormone;
 KW gonadotrophin releasing hormone; fertility control; cancer;
 KW endometriosis; prostate enlargement.

OS Homo sapiens.

PN WO200041720-A1.

PD 20-JUL-2000.

PE 24-DEC-1999; 99WO-AU01167.

PR 08-JAN-1999; 99AU-0008073.

PA (CSLC-) CSL LTD.

PI Walker J;

WPI: 2000-475954/41.

PT Adjuvant composition for manufacturing an immunogenic composition that
 PT can elicit an immune response in an animal, comprises an ionic
 PT polysaccharide component and a saponin component that is an
 PT immunostimulating complex -

PS Disclosure; Page 50; 53pp; English.

CC The present sequence is a peptide fragment of human luteinising hormone
 CC releasing hormone (also known as LHRH, GnRH and gonadotrophin releasing
 CC hormone). It was used to demonstrate the novel adjuvant of the invention,
 CC which has lower reactivity than previous compositions. Vaccination of
 CC humans and animals against LHRH can be used as a method of fertility
 CC control, as well as enabling the control and treatment of disorders of
 CC the reproductive organs, such as testicular, breast, prostate and ovarian
 CC cancers, prostate enlargement and endometriosis. The composition of the
 CC invention contains an anionic macromolecule and a saponin component, the
 CC latter of which is an immunostimulant, and it can also be used with other
 CC immunogens including soluble protein antigens, peptide haptens conjugated
 CC to a carrier protein and whole viruses.

SQ Sequence 9 AA;

Query Match 96.3%; Score 52; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||
 Db 1 HWSYGLRP 8

RESULT 5

AAB08104

ID AAB08104 standard; peptide: 9 AA.

AC AAB08104;

DT 04-DEC-2000 (first entry)

DE Amino acid sequence of truncated luteinising hormone releasing hormone.

KW T helper cell epitope; CDV; immune response; canine vaccine;

KW luteinising hormone releasing hormone; LHRH.

OS Canis sp.

PN WO200046390-A1.

PD 10-AUG-2000.

PE 07-FEB-2000; 2000WO-AU00070.

PR 05-FEB-1999; 99AU-0008533.
 PR 04-AUG-1999; 99AU-0002013.

PA (UYME-) UNIV MELBOURNE.

PA (CSLC-) CSL LTD.

PA (COMMONWEALTH SCI & IND RES ORG.

PA (COON-) COUNCIL QUEENSLAND INST MEDICAL RES.

PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

PI Jackson DC, Sourav G, Walker J;

WPI: 2000-532904/48.

PT Novel T helper cell epitopes derived from canine distemper virus useful
 PT for preparation of canine vaccines -
 PS Example 3; Page 21; 54pp; English.

CC The present sequence represents luteinising hormone releasing hormone
 CC (LHRH). It is used in vaccines with T helper cell epitopes
 CC AAB08076-B08101, derived from canine distemper virus (CDV). Compositions
 CC comprising these T cell helper epitopes are useful for inducing an
 CC immune response in an animal. The epitopes are useful as components
 CC of animal, in particular, canine vaccines, either simply as synthetic
 CC peptide based vaccines and as additions to vaccines containing more
 CC complex antigens.

SQ Sequence 9 AA;

Query Match 96.3%; Score 52; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
 |||||
 Db 1 HWSYGLRP 8

RESULT 6

AAB90972 standard; Peptide: 9 AA.

AC AAB90972;

DT 22-JUN-2001 (first entry)

DE Luteinising hormone releasing hormone (LH-RH) related peptide SEQ ID:146.

KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimidyl; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

OS Homo sapiens.
 OS Synthetic.


```

XX LH-RH, liberating factor for LH and FSH, and its agonists compsn.
PT - used to treat prostate adenocarcinoma, benign hypertrophy of
PT the prostate, hirsutism, acne, etc.
XX
PS Claim 1(d); Page 15; 27pp; French.
XX
CC A composition is claimed containing LHRH or its analogues. The
CC composition is used to treat prostate adenocarcinoma, benign
CC hypertrophy of the prostate, endometriosis, dysmenorrhea, hirsutism,
CC hormone-dependent mammary tumours, for treatment or prevention of
CC precocious puberty, delaying the onset of puberty and for treating
CC acne. The compositions may also contain antiandrogens.
CC See AAP10411-P10418.
CC
XX
SQ Sequence 9 AA:
Query Match 96.3%; Score 52; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9
RESULT 2
AAP50568
ID AAP50568 standard; Protein: 9 AA.
AC AAP50568;
XX
DT 29-NOV-1991 (first entry)
XX
DE Sequence of gonadoliberin analogue IIC.
XX
KW Gonadoliberin agonist; follitropin release; lutropin release;
KW parathormone; hypertension; therapy.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /label= pyroGlu
FT Modified-site 9 /label= bonded to -NHCH3, -NH-CH2-CH3, -NH-CH2CH2CH3
XX
PN DE332329-A.
XX
PD 28-MAR-1985.
XX
PF 08-SEP-1983; 83DE-3332329.
XX
PR 08-SEP-1983; 83DE-3332329.
XX
PA (PARH ) HOECHST AG.
XX
PI Konig W, Neubauer H;
XX
DR WPI; 1985-081717/14.
XX
PT Comprn. contg. gonadoliberin or its analogues - for treating
PT parathormone deficiency states, e.g. hypocalcaemic conditions or
PT hypertension
XX
PS Disclosure: Page 6-7; 17pp; German.
XX
CC The inventors claim a compsn. for treating metabolic disorders
CC caused by inadequate secretion of endogenous parathormone (PTH)
CC which contains, apart from an acceptable carrier, gonadoliberin or
CC agonists at least as strongly active as gonadoliberin. For
CC &parental use these provide 0.5-5 micrograms gonadoliberin per unit
CC dose, and for application to mucosa (intranasally) 10-200 micrograms
CC per dose, for an adult of average wt.
XX

```

```

SQ Sequence 9 AA:
Query Match 96.3%; Score 52; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9
RESULT 3
AAW94891
ID AAW94891 standard; peptide: 9 AA.
AC AAW94891;
XX
DT 11-MAY-1999 (first entry)
XX
DE LHRH peptide fragment.
XX
KW LHRH: immune response; luteinising hormone releasing hormone; DT;
KW diphtheria toxoid; castrating; oestrus cycling; aggression; breast;
KW sexual activity; organoleptic; livestock; cell growth; malignant;
KW prostate; ovarian; oncofoetal; hyperplastic; pregnancy;
KW endometriosis; inflammatory response.
XX
OS Homo sapiens.
XX
PN W09902180-A1.
XX
PD 21-JAN-1999.
XX
PF 09-JUL-1998; 98WO-AU00532.
XX
PR 09-JUL-1997; 97AU-0007768.
XX
PA (CSLC-) CSL LTD.
XX
PI McNamara MK;
XX
DR WPI: 1999-120511/10.
XX
PT New immunogenic luteinising hormone releasing hormone compositions -
PT comprise LHRH conjugated to diphtheria toxoid and adsorbed to an
PT ionic polysaccharide, used to inhibit reproductive function in
PT animals
XX
PS Example 3; Page 30; 41pp; English.
XX
CC The invention relates immunogenic composition for eliciting an immune
CC response to luteinising hormone releasing hormone (LHRH). The
CC composition comprises a LHRH-diphtheria toxoid (DT) conjugate adsorbed to
CC an ionic polysaccharide. The LHRH-DT compositions can be used for
CC eliciting an immune response to LHRH, for castrating an animal, for
CC regulating oestrus cycling in a female animal or for inhibiting
CC characteristics induced by the sexual maturation of an animal, e.g.
CC aggression or sexual activity. They can also be used for achieving
CC production gains in livestock, e.g. reduction or elimination of unwanted
CC organoleptic characteristics from the meat of livestock. They can also be
CC used for inhibiting the growth of cells which are regulated directly or
CC indirectly by LHRH, e.g. malignant breast cells, malignant prostate
CC cells, malignant ovarian cells, malignant oncofoetal cells or
CC hyperplastic cells. They can also be used for down-regulating the libido
CC of an animal. They can also be used for inhibiting pregnancy, prostate
CC enlargement, endometriosis or inflammatory responses. The LHRH
CC compositions induce a more effective immune response against LHRH than
CC the LHRH-carrier-adjutant compositions. The effective immune response
CC against LHRH results in prevention of the release of the hormones LH and
CC FSH from the anterior pituitary. Sequences AAW94890-93 are peptide
CC derivatives of LHRH.
XX
SQ Sequence 9 AA:

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 5.25661 Seconds
(without alignments)
21.311 Million cell updates/sec

Title:	US-09-848-834A-1
Perfect score:	54
Sequence:	1 XHWSYGLRPX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      747574 segs, 111073796 residues
Total number of hits satisfying chosen parameters:  747574
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

1:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1980.DAT *
2:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1981.DAT *
3:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1982.DAT *
4:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1983.DAT *
5:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1984.DAT *
6:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1985.DAT *
7:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1986.DAT *
8:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1987.DAT *
9:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1988.DAT *
10:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1989.DAT *
11:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1990.DAT *
12:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1991.DAT *
13:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1992.DAT *
14:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1993.DAT *
15:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1994.DAT *
16:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1995.DAT *
17:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1996.DAT *
18:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1997.DAT *
19:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1998.DAT *
20:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1999.DAT *
21:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA2000.DAT *
22:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA2001.DAT *

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	52	96.3	9	2	AAPI0414	Luteinising Hormon
2	52	96.3	9	6	AAFP5058	Sequence of gonad
3	52	96.3	9	20	AAW94891	LHRH peptide frag
4	52	96.3	9	21	AAI15363	Human LHRH peptid
5	52	96.3	9	21	AA808104	Amnio acid sequen
6	52	96.3	9	22	AA890972	Luteinising hormo
7	52	96.3	9	22	AA890979	Luteinising hormo
8	52	96.3	9	22	AA859836	GNRH peptide.. pe
9	52	96.3	10	2	AAPI0097	Sequence of lutein
10	52	96.3	10	2	AAPI0411	Luteinising Hormon
11	52	96.3	10	2	AAPI0416	Luteinising Hormon

12	52	96.3	10	3	AAp20277	Modified carboxy t
13	52	96.3	10	6	AAp50222	Gonadotropin rele
14	52	96.3	10	7	AAp60127	Gonadoliberin anta
15	52	96.3	10	7	AAp61403	Gonadotropin relea
16	52	96.3	10	7	AAp60576	Novel decapeptide
17	52	96.3	10	8	AAp70922	Novelinsing hormon
18	52	96.3	10	10	AAp90630	Sequence of lutein
19	52	96.3	10	12	AAr15713	Peptide #1 with ho
20	52	96.3	10	13	AAr26819	LH releasing hormo
21	52	96.3	10	14	AAr33434	Therapeutic agent
22	52	96.3	10	15	AAr62689	LHRH hapten for att
23	52	96.3	10	16	AAr91197	LHRH peptide . SYM
24	52	96.3	10	16	AAr86845	Gonadotropin relea
25	52	96.3	10	16	AAp75152	Gonadotropin relea
26	52	96.3	10	17	AAp65201	Luteinlising hormon
27	52	96.3	10	17	AAp65203	Luteinlising hormon
28	52	96.3	10	18	AAW45642	Luteinlising hormon
29	52	96.3	10	18	AAW22390	Gonadotropin relea
30	52	96.3	10	18	AAW16390	Gonadotropin relea
31	52	96.3	10	18	AAW04612	Gonadotropin relea
32	52	96.3	10	19	AAW79566	Luteinlizing hormone
33	52	96.3	10	19	AAW76381	GnRH-1 polypeptida
34	52	96.3	10	19	AAW76373	Rat modified GnRH
35	52	96.3	10	19	AAW61341	Rat GnRH peptide.
36	52	96.3	10	20	AAW50229	Peptide hormone Gn
37	52	96.3	10	20	AAW31176	Neutrophil-activat
38	52	96.3	10	20	AAW31180	ubiquitin fusion p
39	52	96.3	10	20	AAW31067	ubiquitin fusion p
40	52	96.3	10	20	AAW03864	Non-crosslinked p
41	52	96.3	10	20	AAW03856	Amino acid sequenc
42	52	96.3	10	20	AAW94890	Amino acid sequenc
43	52	96.3	10	20	AAW96765	LHRH peptide fragm
44	52	96.3	10	20	AAW84278	Luteinlising hormon
45	52	96.3	10	20	AAW84286	Hormone domain of
						Modified hormone d

ALIGNMENTS

RESULT 1	
AAPl0414	
ID	AAPl0414 standard; Protein; 9 AA.
XX	
AC	AAPl0414;
XX	
DT	17-DEC-1992 (first entry)
XX	
DE	Luteinising Hormone Releasing Hormone analogue #3.
XX	
XX	LHRH; Follicle Stimulating Factor; FSH; acne; hirsutism;
RW	dysmenorrhoea; precocious puberty; endometriosis; prostat
XX	benign prostate hypertrophy; mammary tumour.
XX	
EH	Key
FT	Modified-site
FT	1
FT	/label= OTHER
FT	/note= "pyroglutamic acid"
FT	Modified-site
FT	9
FT	/note= "Pro-NH-(CH2)n-CH3 (n=0-2),
FT	Pro-NH-(CH2)2-OH or protected by
XX	pyrrolidino or morpholino gp."
XX	
PN	BE885308-A.
XX	
DD	19-MAR-1981.
XX	
PF	23-FEB-1983;
XX	83BE-0468932.
XX	
PR	21-SEP-1979;
XX	79FR-0023545.
XX	
PA	(ROUS) ROUSSEL UCLAF.
XX	
OR	WPI; 1981-23409D/14 (23409D).

BE8885308-A.
19-MAR-1981.
23-FEB-1983; 83BE-0468932.
21-SEP-1979; 79FR-0023545.
(R0US) ROUSSEL UCLAF.
WFI; 1981-23409D/14 (23409D)

Qy	2	HWSYGLRP	9
			1
Db	25	HWSYGLRP	32

Search completed: October 10, 2002, 16:09:52
Job time : 6.88889 secs

Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
DB 25 HWSYGLRP 32

RESULT 13

Q92094

ID Q92094 PRELIMINARY; PRT; 82 AA.

AC Q92094;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOLIBERIN PRECURSOR (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH) (LULIBERIN).
GN PREPRO-GNRH-1.
OS Oncorhynchus nerka (Sockeye salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8023;
[1]

RP SEQUENCE FROM N.A.
RC STRAIN=NIRKO; TISSUE=BRAIN;
RX MEDLINE=96020547; PubMed=8546809;
RA Asahihara M., Suzuki M., Kubokawa K., Yoshiura Y., Kobayashi M.,
RA Urano A., Alda K.;
RT "Two different precursor genes for the salmon-type gonadotropin-releasing hormone exist in salmonids.";
RT J. Mol. Endocrinol. 15:1-9(1995).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC EMBL: D31868; BAA06666.1; -;
DR InterPro: IPR002047; AKH.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH.1.
DR PROSITE: PS00256; AKH; UNKNOWN_1.
DR PROSITE: PS00473; GNRH.1.
KW Amidation; Hormone; Signal.
FT SIGNAL 1 23
FT CHAIN 24 33
FT CHAIN 37 82
SQ SEQUENCE 82 AA; 9126 MW; C64044EA521B28B CRC64;

Query Match 72.2%; Score 39; DB 13; Length 82;
Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
DB 25 HWSYGLRP 32

RESULT 14

Q9W7G1

ID Q9W7G1 PRELIMINARY; PRT; 82 AA.

AC Q9W7G1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH) (LULIBERIN) (GONADOTROPIN-RELEASING HORMONE PRECURSOR I ISOFORM B').
GN GNRH1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
[1]

RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene differs in brain and gonads in rainbow trout.";
RT J. Endocrinology 140:3012-3024(1999).
RN [12]

RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
RN [13]

RX MEDLINE=21232987; PubMed=11335940;
RA Uzbekova S., Ferriture F., Guiguen Y., Bailhache T., Breton B.,
RA Lareyre J.J.;
RT "Stage-dependent and alternative splicing of sGRH messengers in rainbow trout testis during spermatogenesis.";
RT Mol. Reprod. Dev. 59:1-10(2001).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC EMBL: AF110992; AAD3462.1; -;
DR EMBL: AF269108; AAK54680.1; -;
DR InterPro: IPR002047; AKH.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH.1.
DR PROSITE: PS00256; AKH; UNKNOWN_1.
DR PROSITE: PS00473; GNRH.1.
KW Amidation; Hormone.
SQ SEQUENCE 82 AA; 9232 MW; 7595B4FCC65FDFD6 CRC64;

Query Match 72.2%; Score 39; DB 13; Length 82;
Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
DB 25 HWSYGLRP 32

RESULT 15

G90VY3

ID G90VY3 PRELIMINARY; PRT; 82 AA.

AC G90VY3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOTROPIN-RELEASING HORMONE PRECURSOR II ISOFORM B2 (GONADOTROPIN-RELEASING HORMONE PRECURSOR II ISOFORM A)
DE (GONADOTROPIN-RELEASING HORMONE PRECURSOR II ISOFORM B1).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
[1]

RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=21232987; PubMed=11335940;
RA Uzbekova S., Ferriture F., Guiguen Y., Bailhache T., Breton B.,
RA Lareyre J.J.;
RT "Stage-dependent and alternative splicing of sGRH messengers in rainbow trout testis during spermatogenesis.";
RT Mol. Reprod. Dev. 59:1-10(2001).
RL EMBL: AF269105; AAK54677.1; -;
DR EMBL: AF269103; AAK54675.1; -;
DR EMBL: AF269104; AAK54676.1; -;
SQ SEQUENCE 82 AA; 9171 MW; 8053F4F23B115408 CRC64;

Query Match 72.2%; Score 39; DB 13; Length 82;
Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DE BETA-1 POLYPEPTIDE (EC 3.6.1.1) (FRAGMENT).
 OS Beta vulgaris (Sugar beet).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
 OX NCBI_TaxID=3555;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92179265; PubMed=1311852;
 RA Sarafian V., Kim Y., Poole R.J., Rea P.A.;
 RT "Molecular cloning and sequence of cDNA encoding the pyrophosphate-
 energized vacuolar membrane proton pump of Arabidopsis thaliana.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:1775-1779(1992).
 SQ SEQUENCE 24 AA; 2396 MW; CE19F75ADBFD43B CRC64;

Query Match 19.8%; Score 32; DB 10; Length 24;
 Best Local Similarity 60.0%; Pred. No. 8.2e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 SSGPSLQYIK 20
 Db 15 TSGPSLNILK 24

RESULT 7

Q9TWK5 ID Q9TWK5 PRELIMINARY; PRT; 31 AA.
 AC Q9TWK5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PROXIMAL COLLAGEN, COL-P.
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95230211; PubMed=7714453;
 RA Qin X., Waite J.H.;
 RT "Exotic collagen gradients in the byssus of the mussel Mytilus
 edulis.";
 RL J. Exp. Biol. 198:633-644(1995).
 DR InterPro; IPR000087; Collagen.
 SQ SEQUENCE 31 AA; 2648 MW; B1F7708959101A73 CRC64;

Query Match 19.8%; Score 32; DB 5; Length 31;
 Best Local Similarity 71.4%; Pred. No. 1.1e+03;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 PGSSGSPS 15
 Db 15 PGSTGPT 21

RESULT 8

Q9F5S3 ID Q9F5S3 PRELIMINARY; PRT; 30 AA.
 AC Q9F5S3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE D-LACTATE DEHYDROGENASE.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 OC Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92031480; PubMed=1931965;
 RA Bugg T.D., Wright G.D., Dutka-Malen S., Arthur M., Courvalin P.,
 RA Walsh C.T.;
 RT "Molecular basis for vancomycin resistance in Enterococcus faecium

RT BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by
 RT vancomycin resistance proteins VanH and VanA.";
 RL Biochemistry 30:10408-10415(1991).
 DR HSP; P26297; IDLD. 3544 MW; B1696A5C11B2038C CRC64;
 SQ SEQUENCE 30 AA; 3544 MW; B1696A5C11B2038C CRC64;

Query Match 18.8%; Score 30.5; DB 2; Length 30;
 Best Local Similarity 42.1%; Pred. No. 1.8e+03;
 Matches 8; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 3 WSYGLRPGSSGSPSLQYIKA 21
 Db 4 FAYGIRDDEK-PSLEWKA 21

RESULT 9

Q9R5R3 ID Q9R5R3 PRELIMINARY; PRT; 23 AA.
 AC Q9R5R3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GFP-GREEN FLAVOPROTEIN.
 OS Photobacterium leiognathi.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;
 OC Photobacterium.
 OX NCBI_TaxID=658;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92081445; PubMed=1746316;
 RA Raibekas A.A.;
 RT "Green flavoprotein from P. leiognathi: purification, characterization
 and identification as the product of the lux G(N) gene.";
 RL J. Biol. Chem. 266:169-176(1991).
 DR HSP; P09142; INFP.
 SQ SEQUENCE 23 AA; 2806 MW; 5C8BCBABI1FBA388F CRC64;

Query Match 18.5%; Score 30; DB 2; Length 23;
 Best Local Similarity 35.0%; Pred. No. 1.6e+03;
 Matches 7; Conservative 2; Mismatches 5; Indels 6; Gaps 1;

QY 3 WSYGL-----RPGSSGSPSL 16
 Db 4 WNYGVFFLNFYHVGQEPSL 23

RESULT 10

Q9ZEQ3 ID Q9ZEQ3 PRELIMINARY; PRT; 24 AA.
 AC Q9ZEQ3;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PORA PROTEIN (FRAGMENT).
 GN PORA.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=111/95;
 RX MEDLINE=99322353; PubMed=10391880;
 RA Wedege E., Caugant D.A., Musacchio A., Saunders N.B., Zollinger W.D.;
 RT "Redesignation of a purported Pl.15 subtype-specific meningococcal
 monoclonal antibody as a Pl.19-specific reagent.";
 RL Clin. Diagn. Lab. Immunol. 6:639-642(1999).
 DR EMBL; AJ012731; CAA10154.1;
 FT NON_TER 1
 FT NON_TER 24
 SQ SEQUENCE 24 AA; 2670 MW; 260BA0002D72DB7C CRC64;

Query Match 18.5%; Score 30; DB 2; Length 24;
 Best Local Similarity 40.0%; Pred. No. 1.6e+03;

Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 PGSSGSLQYIKANS 23
| : | | : | | |
Db 10 PSNSQPVKTKAKS 24

RESULT 11

Q96H37 ID Q96H37 PRELIMINARY; PRT; 24 AA.
AC Q96H37;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:2676).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008940; AAH08940.1; -;
SQ SEQUENCE 24 AA; 2592 MW; 16AE14B2D2514D01 CRC64;

Query Match 18.5%; Score 30; DB 4; Length 24;
Best Local Similarity 70.0%; Pred. No. 1.6e+03;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 LRPSSGSL 16
| | | | |
Db 10 LPPHSLGPSL 19

RESULT 12

Q9BSM8 ID Q9BSM8 PRELIMINARY; PRT; 26 AA.
AC Q9BSM8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 2.8 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004942; AAH04942.1; -;
KW Hypothetical protein.
SQ SEQUENCE 26 AA; 2808 MW; 59540636D4E0BB9 CRC64;

Query Match 18.5%; Score 30; DB 4; Length 26;
Best Local Similarity 46.7%; Pred. No. 1.8e+03;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 10 GSSGSLQYIKANSK 24
| : | | | | |
Db 5 GOALPSAHVWKANSR 19

RESULT 13

Q93JY7 ID Q93JY7 PRELIMINARY; PRT; 28 AA.
AC Q93JY7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE DNA-DIRECTED RNA POLYMERASE (FRAGMENT).
GN RPOB.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PATIENT NO.30;
RA Cheruvu M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PATIENT NO.30;
RX MEDLINE=21367865; PubMed=11474030;
RA Mani C., Selvakumar N., Narayanan S., Narayanan P.R.;
RT "Mutations in the rpoB gene of multidrug-resistant Mycobacterium tuberculosis clinical isolates from India.";
RL J. Clin. Microbiol. 39:2987-2990(2001).
DR EMBL; AJ27926; CAC50347.1; -;
KW DNA-directed RNA polymerase.
FT NON_TER 1
FT NON_TER 28
SQ SEQUENCE 28 AA; 3071 MW; 678638EFD9DFD446 CRC64;

Query Match 18.5%; Score 30; DB 2; Length 28;
Best Local Similarity 38.5%; Pred. No. 1.9e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 17 OYIKANSKFIGIT 29
| : | | : | | |
Db 7 QFMKQNNPLSGLT 19

RESULT 14

Q45966 ID Q45966 PRELIMINARY; PRT; 30 AA.
AC Q45966;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SCVA PROTEIN.
GN SCVA.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Coxiella group; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NINE MILE;
RX MEDLINE=97055417; PubMed=8899704;
RA Heizen R.A., Howe D., Mallavia L.P., Rockey D.D., Hackstadt T.;
RT "Developmentally regulated synthesis of an unusually small, basic peptide by Coxiella burnetii.";
RL Mol. Microbiol. 22:9-19(1996).
DR EMBL; L49019; AAB39278.1; -;
SQ SEQUENCE 30 AA; 3614 MW; 44D6BBBCA242FC04 CRC64;

Query Match 18.5%; Score 30; DB 2; Length 30;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 RPSGSGP 14
| | | | |
Db 18 RPSGSGP 24

RESULT 15

Q9QUZ3 ID Q9QUZ3 PRELIMINARY; PRT; 15 AA.
AC Q9QUZ3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE BETA-CRYSTALLIN ISOFORM A3/A1 (FRAGMENT).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96030616; PubMed=7556477;
 RA Shearer T.R., Shih M., Azuma M., David L.L.;
 RT "Precipitation of crystallins from young rat lens by endogenous
 RT calpain.";
 RL Exp. Eye Res. 61:141-150(1995).
 SQ SEQUENCE 15 AA; 1702 MW; 3F35688E1C5F233C CRC64;

Query Match 17.9%; Score 29; DB 11; Length 15;
 Best local Similarity 83.3%; Pred. No. 1.4e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 PGSSGP 14
 ||| ||
 Db 4 PGSMGP 9

Search completed: October 10, 2002, 16:48:28
 Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:39:12 : Search time 22.5 Seconds
(without alignments)
153.035 Million cell updates/sec

Title: US-09-848-834A-15

Perfect score: 162

Sequence: 1 XHWSYGLRPGSSGSLQYIKANSRFIGITEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 281105

Minimum DB seq length: 0

Maximum DB seq length: 31

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	45.9	19	21	AA199055
2	75	46.3	25	21	AA192650
3	74	45.7	15	11	AA1906310
4	74	45.7	15	18	AA1935506
5	74	45.7	15	18	AA1911505
6	74	45.7	15	19	AA1967033
7	74	45.7	15	19	AA1971321
8	74	45.7	15	20	AA1940051
9	74	45.7	15	20	AA1967578
10	74	45.7	15	20	AA1973220
11	74	45.7	15	21	AA1945511

12	74	45.7	15	21	AA192637	Tetanus toxoid T c
13	74	45.7	15	21	AA192625	Foreign epitope P2
14	74	45.7	15	21	AA194427	Amino acid sequenc
15	74	45.7	15	21	AA1970300	Clostridium tetani
16	74	45.7	15	21	AA1944763	Tetanus toxoid pro
17	74	45.7	15	22	AA1911763	Clostridium tetani
18	74	45.7	15	22	AA199515	Vaccine related MH
19	74	45.7	15	22	AA195701	Amino acid sequenc
20	74	45.7	15	22	AA1985451	Wild-type TT830 (t
21	74	45.7	15	22	AA1961956	Tetanus Toxoid uni
22	74	45.7	15	22	AA1920143	Tetanus toxin T-ce
23	74	45.7	15	22	AA1968636	HER-2 B cell pepti
24	74	45.7	15	22	AA1946172	Tetanus toxoid TT8
25	74	45.7	15	22	AA1949071	Tetanus toxoid TT
26	74	45.7	16	18	AA1935445	T-cell stimulatory
27	74	45.7	16	20	AA1929705	Clostridium tetani
28	74	45.7	17	15	AA1962692	Helper T cell epit
29	74	45.7	17	16	AA1982573	Tetanus toxin help
30	74	45.7	17	17	AA1905599	Tetanus toxin help
31	74	45.7	17	17	AA1988395	T-cell antigen TT2
32	74	45.7	17	21	AA199274	HLA class II bindi
33	74	45.7	17	21	AA190056	Pathogen derived T
34	74	45.7	17	21	AA194539	T helper cell (Th)
35	74	45.7	17	21	AA1958768	Unidentified pepti
36	74	45.7	17	22	AA199516	Vaccine related MH
37	74	45.7	17	22	AA1962904	Amino acid residu
38	74	45.7	17	22	AA1984435	Amino acid sequenc
39	74	45.7	17	22	AA1930941	Amino acid sequenc
40	74	45.7	17	22	AA1931029	Antigenic fragment
41	74	45.7	17	22	AA1931118	Antigenic fragment
42	74	45.7	17	22	AA1915589	Peptide 5 for pept
43	74	45.7	18	20	AA1926607	HIV-derived lipope
44	74	45.7	19	22	AA199517	Vaccine related MH
45	74	45.7	22	22	AA1946175	Tetanus toxoid 830

ALIGNMENTS

RESULT 1
AA199055
ID AA199055 standard; Peptide; 19 AA.
XX AC
XX AA199055;
XX 07-AUG-2000 (first entry)
XX HLA class II binding antigen epitope peptide #244.
DE Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
XX immune response; chronic viral disease; cancer; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
KW allograft rejection; allergy; Lyme disease; hepatitis; prostate cancer;
KW glomerulonephritis; food hypersensitivity; malaria.
XX Clostridium tetani.
XX WO9961916-A1.
XX 02-DEC-1999.
XX 28-MAY-1999; 99WO-US12066.
XX 29-MAY-1998; 98US-0087192.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Southwood S, Sidney J;
XX WPI; 2000-097143/08.
XX New compositions containing immunogenic peptide epitopes for various
PT HLA class II DR molecules useful for inducing helper T cell response

XX PS Claim 1; Page 44; 60pp; English.

XX CC The present invention relates to a new pharmaceutical composition

CC comprising a unit dose form of a peptide, or analogue, comprising an

CC epitope selected from those represented by peptides AAY98812-Y99339

CC which are derived from various antigens for various human leucocyte

CC antigen class DR molecules, representative of the world wide population.

CC The peptide/analogue binds to an HLA class II molecule at an IC-50 of

CC less than or equal to 1,000 nM. The pharmaceutical can be used to induce

CC a helper T cell response. The pharmaceutical focuses the immune response

CC towards selected determinants and could therefore be used in cases of

CC chronic viral diseases and cancer. Examples of diseases that can be

CC treated using the peptide containing pharmaceutical include autoimmune

CC diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia

CC gravis), allograft rejection, allergies, Lyme disease, hepatitis,

CC post-streptococcal endocarditis or glomerulonephritis and food

CC hypersensitivities. The peptide epitopes can be used to enhance immune

CC responses against other immunogens administered with the peptides.

CC Diseases which can be treated using immunogenic mixtures include

CC prostate cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma,

CC cervical carcinoma, lymphoma, and condyloma acuminatum. The peptides may

CC also be used to make monoclonal antibodies useful as potential diagnostic

CC or therapeutic agents. The peptides may also be useful as diagnostic

CC reagents, for example, to determine the susceptibility of an individual

CC to a treatment regimen. Also, the peptides may be used to predict which

CC individuals will be at substantial risk of developing chronic infection.

CC The selection of appropriate T and B cell epitopes should allow the

CC development of epitope based vaccines particularly towards conserved

CC epitopes of pathogens which are characterized by high sequence

CC variability such as HIV, HCV and Malaria.

XX SQ Sequence 19 AA;

Query Match 46.9%; Score 76; DB 21; Length 19;

Best Local Similarity 93.8%; Pred. No. 0.00022;

Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31

:|||||

Db 2 MOYIKANSKFIGITEL 17

RESULT 2

AAY92650

ID AAY92650 standard; Peptide; 25 AA.

XX AC AAY92650;

XX DT 10-AUG-2000 (first entry)

XX DE PSMpep007 - P2 inserted in hPSM insertion position 6.

XX DE Foreign epitope; P2: prostate specific membrane antigen; vaccination;

KW cytotoxic f-lymphocyte immunity; self-protein; cancer; breast cancer;

KW prostate cancer; cell-associated peptide antigen.

XX OS Synthetic.

OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 6..20

FT /label= P2

FT WO200020027-A2.

PN 13-APR-2000.

XX PD 05-OCT-1999; 99WO-DK00525.

XX PD 05-OCT-1998; 98DK-0001261.

PR 20-OCT-1998; 98US-0105011.

XX DR

(MEBI-) M & E BIOTECH AS.

PA Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

XX Gautam A, Birk P, Karlsson G;

PI WPI; 2000-349917/30.

XX DR

Inducing immune responses to weakly immunogenic, tumor associated

PT peptide antigens for the treatment of breast and prostate cancer

XX Example 1; Page 117; 220pp; English.

XX AAY92650-55 are peptides designed which correspond to the P2 and P30

CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)

CC amino acids in each end. The flanking amino acids correspond to the

CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g.

CC T cell proliferation assays, but also for ELISA or other in vitro

CC assays. The claims detail a method for inducing immune responses against

CC weakly immunogenic cell-associated peptide antigens (PA) such as those

CC associated with cancers (i.e. self-proteins), for example, hPSM,

CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The

CC method comprises effecting simultaneous presentation by antigen

CC producing cells (APCs) of the animals immune system of: (1) at least 1

CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1

CC B-cell group derived from the cell-associated PA; and (2) at least 1

CC first T helper cell group which is foreign to the animal. Analogues of

CC human PSM, human Her2 and human/murine FGF8b comprising a substantial

CC part of all known and predicted CTL and B-cell epitopes of the

CC respective PA and including at least one foreign T helper epitope (e.g.

CC P2 and/or P30) are also claimed. The method is used to treat prostate,

CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and

CC Her2, respectively.

XX SQ Sequence 25 AA;

Query Match 46.3%; Score 75; DB 21; Length 25;

Best Local Similarity 93.8%; Pred. No. 0.00042;

Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31

:|||||

Db 5 VOYIKANSKFIGITEL 20

RESULT 3

AAR06310

ID AAR06310 standard; protein; 15 AA.

XX AC AAR06310;

XX DT 04-DEC-1990 (first entry)

XX DE Tetanus toxin epitope.

XX DE Tetanus toxin; vaccine; major histocompatibility complex; MHC;

KW antimalarial.

XX OS Synthetic.

XX EP378881-A.

XX PD 25-JUL-1990.

XX PF 27-DEC-1989; 89EP-0203318.

XX PR 16-NOV-1989; 89IT-0022409.

XX PR 17-JAN-1989; 89IT-0019110.

XX PA (ENIE) ENRICERCHE SPA.

XX PI Pessi A, Bianchi E, Verdini AS, Corradin G;

XX WPI; 1990-225582/30.

XX Synthetic peptide(s) corresp. to tetanus toxin epitope(s) - used
PT as universal carriers for prepn. of immunogenic conjugate(s) for
PT use as vaccines.
XX
PS Claim 1; Page 17; 20pp; English.
XX
CC Epitopic peptides may be used with synthetic hapten derived from
CC a pathogen to generate an immune response to the pathogen.
CC Peptides are recognised by numerous T-helper cell clones within
CC the context of a wide range of alleles of the human MHC.
CC The peptides may be used in an antimalarial vaccine inducing Ab.
CC response to P.falciparum.
XX
SQ Sequence 15 AA;
Query Match 45.7%; Score 74; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 QYIKANSKFIGITEL 31
DB 1 QYIKANSKFIGITEL 15
IIIIIIIIIIIIIIIIIIII
RESULT 4
AAW35506
ID AAW35506 standard; peptide; 15 AA.
XX
AC AAW35506;
XX
DT 22-APR-1998 (first entry)
XX
DE Universal T-cell epitope peptide SEQ ID NO:8.
XX
KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
KW scaffold; inhibition; metastasis; wound healing; solid phase.
XX
OS Unidentified.
XX
PN WO9738011-A1.
XX
PD 16-OCT-1997.
XX
PF 03-APR-1997; 97WO-DE00146.
XX
PR 03-APR-1996; 96DK-0000398.
XX
PA (PEPR-) PEPRESEARCH AS.
XX
PI Heegaard PMH, Jakobsen PH;
XX
DR WPI; 1997-512645/47.
XX
PT Non-dendritic peptide carrier linked to a solid phase - useful as a
PT diagnostic agent and as a scaffold for production of chemical
PT derivatives
XX
PS Example 20; Page 124; 262pp; English.
XX
CC A non-dendritic peptide carrier (A) has been developed which is coupled
CC through a linker to a solid phase, forming a complex of (A)-solid phase.
CC Where (A) comprises 10-50 amino acids capable of forming a secondary
CC structure in a benign buffer after liberation from the solid phase, and
CC further the (A)-solid phase complex comprises an immunogenic substance
CC and/or an immune mediator coupled on (A). The present sequence
CC represents a peptide used in an example from the present invention.
CC (A)-solid phase complex can be used as a scaffold for the production of
CC chemical derivatives, characterised by covalently attaching molecules at
CC attachment points. Alternatively (A) is used as a scaffold-peptide for
CC the incorporation into an immunostimulating complex (iscom) resulting an
CC (A)-iscom complex which is used for the chemical coupling of antigenic
CC substances in an aqueous solution by conjugation. (A) derivatised with

CC one or more peptides having fibronectin-, laminin- or vitronectin-like
CC binding activities can be used for the promotion of cell-attachment to
CC plastic surfaces, in particular to inhibit tumour growth and metastasis,
CC and for promotion of wound healing. Also a derivatised (A) can be used
CC for the selection of specifically-binding aptamers or as a diagnostic
CC agent. Such diagnostic-(A) molecules could be used to detect molecules
CC derived from or indicative of pregnancy or of a disease, such as an
CC infectious, autoimmune or cancerous disease.
XX
SQ Sequence 15 AA;
Query Match 45.7%; Score 74; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 QYIKANSKFIGITEL 31
DB 1 QYIKANSKFIGITEL 15
IIIIIIIIIIIIIIIIIIII
RESULT 5
AAW11505
ID AAW11505 standard; Protein; 15 AA.
XX
AC AAW11505;
XX
DT 24-SEP-1997 (first entry)
XX
DE Tetanus toxoid universal Th epitope TT830.
XX
KW Humanised antibody; anti-Fc receptor; H22; bifunctional; bispecific;
KW fusion protein; chimera; tetanus toxoid; helper T cell epitope;
KW antigen presentation; ds.
XX
OS Clostridium tetani.
XX
PN WO9640789-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09988.
XX
PR 07-JUN-1995; 95US-0484172.
XX
PA (MEDA-) MEDAREX INC.
XX
PI Deo YM, Goldstein J, Graziano R, Somasundaram C;
XX
DR WPI; 1997-052242/05.
XX
DR N-PSDB; AAT58127.
XX
PT Recombinant, multi-specific anti-Fc receptor antibody molecules -
PT also comprise an anti-target portion, used for the treatment of
PT cancer, autoimmune disease and pathogenic infection
XX
PS Example 7; Fig 24; 115pp; English.
XX
CC Synthetic DNA coding for the wild-type universal Th epitope from
CC tetanus toxoid, designated TT830, was fused to the 3'-end of DNA
CC encoding heavy chain sequences from the humanised anti-Fc gamma RI
CC monoclonal antibody H22. The resulting fusion protein was shown to
CC be significantly more efficient in antigen presentation and T cell
CC stimulation than the TT830 epitope alone. A similar fusion
CC construct was prepared coding for a mutant, antagonistic form of the
CC epitope (designated TT833S) fused to the anti-Fc gamma RI. The
CC Fc22-TT833S is at least 100 times more effective than TT833S in
CC inhibiting T cell activation.
XX
SQ Sequence 15 AA;
Query Match 45.7%; Score 74; DB 18; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
 DB 1 QYIKANSKFIGITEL 15

RESULT 6
 AAW67033
 ID AAW67033 standard; peptide; 15 AA.
 AC AAW67033;
 XX 15-DEC-1998 (first entry)
 DT Tetanus toxin fragment (residues 830-844).
 DE Tetanus toxin; vaccine; antibody; carbohydrate peptide conjugate;
 XX dendrimeric poly-lysine; epitope; tumour.
 KW Clostridium tetani.
 OS WO9843677-A1.
 PN 08-OCT-1998.
 XX 27-MAR-1998; 98WO-EP01922.
 XX 27-MAR-1997; 97US-0041726.
 XX (INSP) INST PASTEUR.
 PA Bay S, Cantacuzene D, Leclerc C, Lo-man R;
 PI WPI; 1998-557071/47.
 XX Carbohydrate peptide conjugate used as vaccine - comprises carrier
 PT with dendrimeric poly-lysine enabling multiple epitopes to be
 PT covalently attached
 XX Disclosure; Page 13; 55pp; English.
 XX The invention relates to a new carbohydrate peptide conjugate, which
 CC comprises a carrier with a dendrimeric poly-lysine enabling multiple
 CC epitopes to be covalently attached to it. Also claimed are: (1) an
 CC antibody purified from biological fluid or cells of organisms
 CC administered with the carbohydrate peptide conjugate, and (2) a diagnosis
 CC kit comprising antigen-specific antibodies elicited by immunisation with
 CC the carbohydrate peptide conjugate. The peptide conjugate, antibody and
 CC diagnosis kit are used to provide pharmaceutical compositions and
 CC vaccines against tumours. These can be used to support an immune response
 CC against viral infections caused by hepatitis virus, HIV or cytomegalo
 CC virus. They can be used to enhance immune responses, especially B- and T-
 CC cell responses, of humans and animals against bacterial infections. The
 CC carbohydrate peptide conjugate stimulates the antibody and T-cell
 CC response without stimulating undesired immune responses. The composition
 CC is capable of increasing the survival of tumour bearing humans and
 CC animals. The present sequence corresponds to residues 830-844 of tetanus
 CC toxin. The synthetic peptide corresponding to this sequence may be used
 CC as an epitope in a carbohydrate peptide conjugate.
 XX Sequence 15 AA;
 SQ

Query Match 45.7%; Score 74; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00032;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
 DB 1 QYIKANSKFIGITEL 15

RESULT 7
 AAW71321

ID AAW71321 standard; peptide; 15 AA.
 XX AAW71321;
 AC 26-NOV-1998 (first entry)
 DT Universal helper T-cell epitope P2 derived from tetanus toxin.
 DE Liver stage; Plasmodium; Navy Yoelii Liver Stage 3 antibody; NYLS3;
 XX hepatic and erythrocytic stage protein; PyHEP17; vaccine;
 KW malaria parasite; teanus toxin; P2; helper T-cell epitope.
 XX Synthetic.
 OS Clostridium tetani.
 XX US5814617-A.
 PN 29-SEP-1998.
 XX 07-OCT-1994; 94US-0319704.
 XX 07-OCT-1994; 94US-0319704.
 XX (USNA) US SEC OF NAVY.
 XX Charoenvit Y, Doolan DL, Hedstrom RC, Hoffman SL;
 XX WPI; 1998-541794/46.
 XX Vaccine for protecting mammal against infection by malaria caused by
 PT plasmodium species - comprises a first nucleic acid encoding a first
 PT polypeptide capable of eliciting an immune reaction against an
 PT antigen expressed during the liver
 XX Disclosure; Column 12; 24pp; English.
 XX AAW71321-22 represent universal helper T-cell epitopes derived from
 CC tetanus toxin. They are used to enhance host immune response to
 CC vaccines. The specification describes a Plasmodium yoelii liver stage
 CC 17 kDa hepatic and erythrocytic stage protein designated PyHEP17. This
 CC protein elicits a response from an Ig1 monoclonal antibody designated
 CC Navy Yoelii Liver Stage 3 (NYLS3). This antibody does not recognise
 CC sporozoites, but does recognise P. yoelii liver stage parasites. NYLS3
 CC eliminates upto 90% of liver stage parasites. The specification
 CC describes a vaccine for reducing the severity or incidence of infection
 CC by a malaria parasite of the genus Plasmodium. The DNA vaccine comprises
 CC exon 1 and part of exon 2 of the PyHEP17 gene.
 XX Sequence 15 AA;
 SQ

Query Match 45.7%; Score 74; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00032;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
 DB 1 QYIKANSKFIGITEL 15

RESULT 8
 AAY04051
 ID AAY04051 standard; peptide; 15 AA.
 XX AAY04051;
 AC 04-JAN-2000 (first entry)
 DT T-Helper epitope from tetanus toxoid.
 DE Covalently reactive antigen analog; CRAA; catalytic antibody;
 XX electrophilic reaction centre; phosphonate; boronate; vaccine;
 KW transition state analog; TSA; isostere; gpi20; HIV-1; T-helper;
 KW tetanus; toxoid; B-T-epitope.

XX OS Clostridium tetani.
 XX XX WO9948925-A1.
 PN 30-SEP-1999.
 PD 23-MAR-1999; 99WO-US06325.
 XX 23-MAR-1998; 98US-0046373.
 PR (UYNE-) UNIV NEBRASKA.
 XX Paul S. Gololobov G, Smith L;
 PI WPT; 1999-591076/50.
 DR
 XX New covalently reactive antigen analogs used for treating e.g.
 PT autoimmune diseases, lymphoproliferative disorders, cancers, microbial
 PT infections, ischemic and reperfusion injury or septic shock -
 XX
 PS Disclosure; Page 86; 158pp; English.
 XX The patent discloses new covalently reactive antigen analogs (CRAA)
 CC of formula XI-Y-E-X2, in which X1 and X2 represent peptide sequences
 CC of an epitope of a disease-associated protein, Y is a positively
 CC charged amino acid residue, preferably Lys or Arg, and E is an
 CC electrophilic reaction centre, preferably a phosphonate or boronate
 CC moiety. Depending on the identity of the epitope, the CRAA may be used
 CC to stimulate production of catalytic antibodies specific for
 CC predetermined antigens associated with particular medical disorders.
 CC They may also be used to permanently inactivate endogenously produced
 CC catalytic antibodies produced in certain autoimmune diseases as well as
 CC in certain lymphoproliferative disorders.
 CC Amongst the specifically exemplified CRAAs is one based on residues
 CC 421-436 of a B-cell epitope of gp120 (see AAY04046) which may be used
 CC to counter HIV-1 infections. When used as an immunogen, preferably this
 CC CRAA is conjugated at its N-terminal to a T-helper epitope from tetanus
 CC toxoid. The present sequence represents the T-helper epitope and
 CC corresponds to residues 830-844 of the toxoid.
 XX
 SQ Sequence 15 AA;
 Query Match 45.7%; Score 74; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00032;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 QYIKANSKFIGITEL 31
 DB 1 QYIKANSKFIGITEL 15
 RESULT 9
 AAW67578
 ID AAW67578 standard; peptide; 15 AA.
 AC AAW67578;
 XX
 DT 02-MAR-1999 (first entry)
 DE T-cell epitope peptide #4 for chimeric fimbria/T-cell epitope peptide.
 XX Chimeric; non-typable Haemophilus influenzae; fimbria; T-cell epitope;
 KW immunogenic composition; immune response.
 XX Synthetic.
 OS US5843464-A.
 XX
 PN 01-DEC-1998.
 XX 02-JUN-1995; 95US-0460502.
 XX

PR 02-JUN-1995; 95US-0460502.
 XX (OHIS) UNIV OHIO STATE.
 XX Bakaletz LO, Kaunaya PTP;
 PI WPT; 1999-044514/04.
 XX Synthetic chimeric fimbria peptide - useful for vaccination against
 PT non-typable Haemophilus influenzae
 XX Disclosure; Column 4; 16pp; English.
 XX The invention relates to the manufacture of a synthetic chimeric peptide
 CC comprising a non-typable Haemophilus influenzae fimbria peptide fused via
 CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is
 CC used in immunogenic compositions which induce an immune response against
 CC non-typable Haemophilus influenzae. This sequence represents an example
 CC of a T-cell epitope peptide used to generate the chimeric peptide.
 XX
 SQ Sequence 15 AA;
 Query Match 45.7%; Score 74; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00032;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 QYIKANSKFIGITEL 31
 DB 1 QYIKANSKFIGITEL 15
 RESULT 10
 AAW73220
 ID AAW73220 standard; Protein; 15 AA.
 XX
 AC AAW73220;
 XX
 DT 25-JAN-1999 (first entry)
 DE Tetanus toxoid epitope.
 XX
 KW Multispecific single chain antibody; antibody H22; tumour cell; therapy;
 KW antibody-dependent cellular cytotoxicity; ADCC; HER 2/neu; infection;
 XX epidermal growth factor receptor; breast cancer; ovarian cancer.
 OS Synthetic.
 XX US5837243-A.
 XX 17-NOV-1998.
 XX
 PF 07-JUN-1996; 96US-0661052.
 XX
 PR 07-JUN-1996; 96US-0661052.
 PR 07-JUN-1995; 95US-0484172.
 XX (MEDA-) MEDAREX INC.
 XX
 XX Deo YM, Goldstein J, Graziano R, Somasundaram C;
 PI WPT; 1999-023374/02.
 XX Specific killing of tumour cells - using a multi-specific molecule
 PT comprising an anti-TC receptor antibody and a portion which binds to
 PT a target cell
 XX
 PS Example 7; Column 27; 57pp; English.
 XX This sequence represents a tetanus toxoid epitope and is recognised
 CC by the multispecific single chain antibody designated H22. The
 CC antibody can be used in the method of the invention for inducing
 CC antibody-dependent cellular cytotoxicity (ADCC) against a tumour cell
 CC which is characterised by overexpression of HER 2/neu or epidermal growth

CC factor receptor (EGFR), comprises contacting the tumour cell with a
CC multispecific protein molecule (preferably a single chain antibody)
CC comprising: (a) an anti-Fc receptor antibody or an antigen binding
CC fragment; (b) a portion which binds to HER 2/neu; and (c) a portion which
CC binds to EGFR. The method can be used for treating cancers especially
CC breast cancer or ovarian cancer. The multispecific antibody can also
CC be administered prophylactically to vaccinate a subject against infection
CC by a target cell.
XX
SQ Sequence 15 AA;

Query Match 45.7%; Score 74; DB 20; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
Db 1 QYIKANSKFIGITEL 15

RESULT 11

AAB4511
ID AAB4511 standard; Protein; 15 AA.

XX
AC AAB4511;

XX
DT 26-FEB-2001 (first entry)

XX
DE Tetanus P2 epitope SEQ ID NO: 23.

XX
KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.

XX
OS Clostridium tetani.

XX
PN WO200065058-A1.

XX
PD 02-NOV-2000.

XX
PF 19-APR-2000; 2000WO-DK00205.

XX
PR 23-APR-1999; 99DK-0000552.

XX
PR 06-MAY-1999; 99US-0132811.

XX
PA (MEBI-) M & E BIOTECH AS.

XX
PI Klysner S;

XX
DR WPI; 2000-672791/65.

XX
PT Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -

XX
PS Example 1; Page 137; 172pp; English.

XX
CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.
XX
SQ Sequence 15 AA;

Query Match 45.7%; Score 74; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31

Db 1 QYIKANSKFIGITEL 15

RESULT 12

AAY82637
ID AAY82637 standard; peptide; 15 AA.

XX
AC AAY82637;

XX
DT 07-AUG-2000 (first entry)

XX
DE Tetanus toxoid T cell epitope peptide SEQ ID NO:13.

XX
KW T cell epitope: B cell epitope; allergy; allergen; antigenic;
KW antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW atopic dermatitis; acute urticaria; chronic urticaria;
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.

XX
OS Clostridium tetani.

OS Synthetic.

XX
PN WO200006694-A2.

XX
PD 10-FEB-2000.

XX
PF 20-JUL-1999; 99WO-BE00092.

XX
PR 30-JUL-1998; 98EP-0870167.

XX
PA (UNIO) UCB SA.

XX
PI Saint-Remy J, Jacquemin M;

XX
DR WPI; 2000-422470/36.

XX
PT New compound for prevention and treatment of allergies comprises at
PT least one allergen antigenic determinant recognized by a B cell and at
PT least one antigenic determinant which does not trigger T cell
PT activation -

XX
PS Example 6; Page 30; 50pp; English.

XX
CC The present invention describes a compound (I) for the prevention and/or
CC treatment of allergy. The compound comprises at least one allergen
CC antigenic determinant (i) recognised by a B cell or an antibody secreted
CC by a B cell of a non-atopic individual and at least one antigenic
CC determinant (ii) different from the allergen that triggers T cell
CC activation. (I) has antiallergic, antiasthmatic, antiinflammatory,
CC dermatological and immunosuppressive activities, and can be used in a
CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
CC treat and/or prevent allergies or a disease of allergic origin,
CC especially hypersensitivities. These include rhinitis, sinusitis,
CC bronchial asthma, atopic dermatitis, some forms of acute and chronic
CC urticaria, gastro-intestinal syndromes associated with the ingestion of
CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions
CC associated with drug hypersensitivities and/or a mixture of these. The
CC use of (I) in the treatment of allergic conditions avoids the need for
CC drug treatment, which often causes undesirable side-effects. Also, prior
CC art drug therapies alleviate symptoms, but do not influence their
CC causes, however (I) actually combats the cause of an allergic reaction.
CC The present sequence represents a peptide, which is used in an
CC example from the present invention.
XX
SQ Sequence 15 AA;

Query Match 45.7%; Score 74; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31

Db |||||
-1 QYIKANSKFIGITEL 15

RESULT 13

AA92625
ID AAY92625 standard; Protein; 15 AA.

XX
AC AAY92625;
XX
DT 10-AUG-2000 (first entry)
XX
DE Foreign epitope P2.
XX
KW Foreign epitope; P2; prostate specific membrane antigen; PSM; Her2;
KW Heregulin 2; Fibroblast growth factor 8b; FGF8b; vaccination;
KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW prostate cancer; cell-associated peptide antigen.

XX Clostridium tetani.

OS WO200020027-A2.

PN 13-APR-2000.

PF 05-OCT-1999; 99WO-DK00525.

PR 05-OCT-1998; 98DK-0001261.

PR 20-OCT-1998; 98US-0105011.

XX (MEBI-) M & E BIOTECH AS.

PA Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
PI
XX

DR WPI; 2000-349917/30.
DR N-PSDB; AAA09460.

XX Inducing immune responses to weakly immunogenic, tumor associated
XX peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page 213; 220pp; English.

XX The claims detail a method for inducing immune responses against weakly
XX immunogenic cell-associated peptide antigens (PA) such as those
XX associated with cancers (i.e. self-proteins), for example, human
XX prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
XX fibroblast growth factor 8b (FGF8b). The method comprises effecting
XX simultaneous presentation by antigen producing cells (APCs) of the
XX animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
XX group derived from the PA and/or at least 1 B-cell group derived from the
XX cell-associated PA; and (2) at least 1 first T helper cell group which is
XX foreign to the animal. Analogues of human PSM, human Her2 and
XX human/murine FGF8b comprising a substantial part of all known and
XX predicted CTL and B-cell epitopes of the respective PA and including at
XX least one foreign T helper epitope (e.g. P2 and/or P30) are also claimed.
XX The method is used to treat prostate, prostate/breast or breast cancer
XX when the PA is human PSM, FGF8b and Her2, respectively.

XX Sequence 15 AA;

Query Match 45.7%; Score 74; DB 21; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.00032;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31

Db 1 QYIKANSKFIGITEL 15

RESULT 14

AA984427
ID AAY84427 standard; Peptide; 15 AA.

XX
AC AAY84427;

DT 25-JUL-2000 (first entry)

XX Amino acid sequence of the tetanus toxoid P2 epitope.

DE Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;
KW tumour necrosis factor receptor; type II transmembrane protein;
KW osteoclast differentiation; CSF-1; osteoclast activator;
KW immune response; osteoporosis; bone resorption;
KW tetanus toxoid P2 epitope.

XX Clostridium tetani.

OS WO200015807-A1.

PN 23-MAR-2000.

PF 13-SEP-1999; 99WO-DK00481.

PR 15-SEP-1998; 98DK-0001164.

PR 02-OCT-1998; 98US-0102896.

XX (MEBI-) M & E BIOTECH AS.

XX Halkier T, Haaning J;
PI
XX

DR WPI; 2000-271444/23.

XX In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used
XX to treat, prevent and ameliorate osteoporosis -
XX Example; Page 106; 110pp; English.

XX The present sequence represents the tetanus toxoid P2 epitope. It is
XX used to create a fusion protein with murine osteoprotegerin ligand
XX (OPGL). Osteoprotegerin is a secreted member of the tumour necrosis
XX factor receptor family, which blocks osteoclastogenesis in a dose
XX dependent manner. The OPGL protein is synthesised as a type II
XX transmembrane protein. The murine and human OPGL polypeptides are 87%
XX homologous. OPGL is a potent osteoclast differentiating factor when
XX combined with CSF-1. It is not capable of inducing osteoclast
XX differentiation in the absence of CSF-1. OPGL is also an activator of
XX mature osteoclasts. The specification describes a method for the in vivo
XX down-regulation of OPGL activity in an animal. The method comprises
XX using at least one OPGL polypeptide or subsequence, and/or at least one
XX OPGL analogue to induce an immune response in the animal. The method
XX and OPGL polypeptide are useful for treating, preventing and ameliorating
XX osteoporosis or other diseases or conditions characterised by excessive
XX bone resorption.

XX Sequence 15 AA;

Query Match 45.7%; Score 74; DB 21; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.00032;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31

Db 1 QYIKANSKFIGITEL 15

RESULT 15

AA970300
ID AAY70300 standard; peptide; 15 AA.

XX
AC AAY70300;

XX 06-JUN-2000 (first entry)

XX Clostridium tetani tetanus toxoid T-cell epitope, P589.

KW Recombinant protein; CDC/NIIMALVAC-1; multivalent; malaria; vaccine;
 KW T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
 KW circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
 KW liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
 KW apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
 KW EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
 KW Pf27; antiparasitic; prevention; anti-CDC/NIIMALVAC-1 antibody.
 XX

OS Clostridium tetani.

PN WO200011179-A1.

XX 02-MAR-2000.

PD 19-AUG-1999; 99WO-US18869.

XX 21-AUG-1998; 98US-0097703.

PR (NAIM-) NAT INST IMMUNOLOGY.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Lal AA, Shi YP, Hasnain SE;

XX WPI; 2000-237654/20.

XX Novel recombinant protein as vaccine for treating malarial infection
 PT comprises antigenic peptides obtained from different stages of
 PT plasmodium falciparum life cycle -
 XX

PS Claim 2; Page 17; 52pp; English.

XX The present sequence is the tetanus toxoid T-cell epitope P589, derived
 CC from Clostridium tetani. It is used in the construction of recombinant
 CC protein CDC/NIIMALVAC-1, which is a multivalent, multistage malarial
 CC vaccine. The recombinant protein comprises, melittin signal peptide,
 CC (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
 CC from circumsporozoite protein (CSP), sporozoite surface protein-2
 CC (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 CC (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
 CC antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 CC specific antigen, Pf27. These epitopes were obtained at different stages
 CC of the life cycle of P. falciparum. CDC/NIIMALVAC-1 vaccine has
 CC antiparasitic activity and can be used for treatment and prevention of
 CC malarial infections. Anti-CDC/NIIMALVAC-1 antibodies can be used for
 CC detecting P. falciparum in biological samples.
 XX

SQ Sequence 15 AA;

Query Match 45.7%; Score 74; DB 21; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.00032;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31

Db 1 QYIKANSKFIGITEL 15

Search completed: October 10, 2002, 16:44:06

Job time : 22.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:47:38 ; Search time 11.5 seconds
(without alignments)
65.843 Million cell updates/sec

Title: US-09-848-834A-15

Perfect score: 162

Sequence: 1 XHWSYGLRSGSGPSLOYIRANSKFIGITEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 145535

Minimum DB seq length: 0

Maximum DB seq length: 31

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	% Match	Length	ID	Description
1	84	51.9	24	5	PCT-US92-07218-31
2	76	46.9	17	5	PCT-US95-08596-23
3	76	46.9	31	5	PCT-US93-11703-63
4	74	45.7	15	2	US-08-319-704-10
5	74	45.7	15	2	US-08-661-052-6
6	74	45.7	15	2	US-08-460-502-7
7	74	45.7	15	4	US-09-046-373-2
8	74	45.7	15	4	US-09-188-082-6
9	74	45.7	15	5	PCT-US93-11703-69
10	74	45.7	16	4	US-09-248-588-55
11	74	45.7	17	1	US-08-446-692-4
12	74	45.7	17	2	US-08-488-351A-4
13	74	45.7	17	3	US-09-100-409A-40
14	74	45.7	17	5	PCT-US95-13841-7
15	74	45.7	27	1	US-08-446-692-13
16	74	45.7	27	2	US-08-488-351A-13
17	72	44.4	27	5	PCT-US92-07218-32
18	72	44.4	29	3	US-09-075-257A-13
19	72	44.4	29	3	US-09-075-257A-14
20	72	44.4	29	4	US-09-534-639-13
21	72	44.4	29	4	US-09-534-639-14
22	72	44.4	30	5	PCT-US92-07218-29
23	71	43.8	17	1	US-08-188-223-6
24	71	43.8	17	4	US-08-968-466-6
25	71	43.8	17	4	US-08-478-546B-6
26	70	43.2	14	1	US-08-186-266-5
27	70	43.2	14	1	US-08-305-871A-5

28	70	43.2	14	1	US-08-465-167A-18	Sequence 18, Appl
29	70	43.2	14	2	US-08-817-933A-9	Sequence 9, Appl
30	70	43.2	14	5	PCT-US92-07218-15	Sequence 15, Appl
31	70	43.2	14	5	PCT-US92-07218-30	Sequence 30, Appl
32	70	43.2	14	5	PCT-US95-02121-95	Sequence 95, Appl
33	70	43.2	19	1	US-08-787-547-41	Sequence 41, Appl
34	70	43.2	24	5	PCT-US92-07218-25	Sequence 25, Appl
35	70	43.2	24	5	PCT-US95-02121-110	Sequence 110, App
36	70	43.2	27	5	PCT-US92-07218-26	Sequence 26, Appl
37	70	43.2	27	5	PCT-US92-07218-27	Sequence 27, Appl
38	70	43.2	27	5	PCT-US92-07218-28	Sequence 28, Appl
39	70	43.2	27	5	PCT-US95-02121-111	Sequence 111, App
40	70	43.2	27	5	PCT-US95-02121-112	Sequence 112, App
41	69	42.6	15	2	US-08-661-052-9	Sequence 9, Appl
42	69	42.6	15	4	US-09-188-082-9	Sequence 9, Appl
43	66	40.7	14	4	US-09-082-279B-510	Sequence 510, App
44	66	40.7	14	4	US-09-315-304B-510	Sequence 510, App
45	65	40.1	13	1	US-08-787-547-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
PCT-US92-07218-31
; Sequence 31, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitellelo, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/07218
; FILING DATE: 19920826
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14137-26-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-07218-31

```

RESULT 3
PCT-US93-11703-63
; Sequence 63, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
;

```

```

1 ZIP: 20889-5606
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC Compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10
11 SOFTWARE: PatentIn Release #1.0, Version #1.25
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/08/319,704
16
17 FILING DATE: 07-OCT-1994
18
19 CLASSIFICATION: 424
20
21 ATTORNEY/AGENT INFORMATION:
22
23 NAME: A. David Spevack
24
25 REGISTRATION NUMBER: 24,743
26
27 REFERENCE/DOCKET NUMBER: 75,206
28
29 TELECOMMUNICATION INFORMATION:
30

```

TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-1022
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-319-704-10

Query Match 45.7%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QVIKANSKFIGITEL 31
| | | | | | | | | | | | | | | | |
DB 1 QVIKANSKFIGITEL 15

RESULT 5
US-08-661-052-6
Sequence 6, Application US/08661052
Patent No. 5837243

GENERAL INFORMATION:
APPLICANT: Yashwant M. Deo
APPLICANT: Joel Goldstein
APPLICANT: Robert Graziano
APPLICANT: Chezian Somasundaram
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
TITLE OF INVENTION: OF ANTI-PC RECEPTOR ANTIBODIES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/661,052
FILING DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/484,172
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MX1-043CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-661-052-6

Query Match 45.7%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QVIKANSKFIGITEL 31
| | | | | | | | | | | | | | | | |
DB 1 QVIKANSKFIGITEL 15

RESULT 6
US-08-460-502-7
Sequence 7, Application US/08460502
Patent No. 5843464
GENERAL INFORMATION:
APPLICANT: Bakaletz, Lauren O.
APPLICANT: Kaumaya, Parvin T.
TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter and Griswold
STREET: 800 Superior Avenue
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,502
FILING DATE:

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Golrick, Mary E.
REGISTRATION NUMBER: 34,829
REFERENCE/DOCKET NUMBER: 22727/00120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8458
TELEFAX: (216) 241-0816
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-502-7

Query Match 45.7%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QVIKANSKFIGITEL 31
| | | | | | | | | | | | | | | | |
DB 1 QVIKANSKFIGITEL 15

RESULT 7
US-09-046-373-2
Sequence 2, Application US/09046373
Patent No. 6235714

GENERAL INFORMATION:
APPLICANT: Sudhir Paul
APPLICANT: Larry J. Smith
APPLICANT: Gennady Gololobov
TITLE OF INVENTION: Methods for Identifying Inducers and
TITLE OF INVENTION: Inhibitors of Catalytic Antibodies, Compositions and Their
TITLE OF INVENTION: Use
FILE REFERENCE: UNMC 63123
CURRENT APPLICATION NUMBER: US/09/046,373
CURRENT FILING DATE: 1998-03-23
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 15
TYPE: PRT
ORGANISM: Clostridium tetani
US-09-046-373-2

Query Match 45.7%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
| | | | | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 8

US-09-188-082-6
; Sequence 6, Application US/09188082
; Patent No. 6270765
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chезiаn Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,082
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/661,052
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal

US-09-188-082-6

Query Match 45.7%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
| | | | | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 9

PCT-US93-11703-69
; Sequence 69, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green

; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-11703-69

PCT-US93-11703-69

Query Match 45.7%; Score 74; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
| | | | | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 10

US-09-248-588-55
; Sequence 55, Application US/09248588
; Patent No. 6231864
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and
; their Derivatives
; FILE REFERENCE: SYN-101 4564/69529
; CURRENT APPLICATION NUMBER: US/09/248,588
; CURRENT FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: 60/074537
; EARLIER FILING DATE: 1998-02-12
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Clostridium tetani
; PUBLICATION INFORMATION:
; JOURNAL: Vaccine
; VOLUME: 15
; ISSUE: 4
; PAGES: 377-
; DATE: 1997
; US-09-248-588-55

Query Match 45.7%; Score 74; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
| | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15

RESULT 11

US-08-446-692-4
; Sequence 4, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-446-692-4

Query Match 45.7%; Score 74; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
| | | | | | | | | | | | | | |
Db 3 QYIKANSKFIGITEL 17

RESULT 12

US-08-488-351A-4
; Sequence 4, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US

; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992

; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745

; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-08-488-351A-4

Query Match 45.7%; Score 74; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
| | | | | | | | | | | | | | |
Db 3 QYIKANSKFIGITEL 17

RESULT 13

US-09-100-409A-40
; Sequence 40, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY

; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:

; CLASSIFICATION: 514

```
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-40

Query Match 45.7%; Score 74; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
Db 3 QYIKANSKFIGITEL 17

RESULT 14
PCT-US95-13841-7
; Sequence 7, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEEX: 421792
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-7

Query Match 45.7%; Score 74; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 17 QYIKANSKFIGITEL 31
Db 3 QYIKANSKFIGITEL 17

RESULT 15
US-08-446-692-13
; Sequence 13, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-13

Query Match 45.7%; Score 74; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
Db 3 QYIKANSKFIGITEL 17

Search completed: October 10, 2002, 16:51:26
Job time : 12 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:37:54 ; Search time 16 Seconds
(without alignments)
216.201 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHWSYGLRPGSSGSLDEKTKIAKMKASSVFNVNS 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 28138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 8036

Minimum DB seq length: 0

Maximum DB seq length: 36

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	31.7	10	1 RHPGG	gonadoliberin - pi
2	58	31.7	10	1 RHSHG	gonadoliberin - sh
3	54	29.5	10	1 RHAQ1	gonadoliberin I
4	45	24.6	10	2 A21114	gonadoliberin - ch
5	40	21.9	10	1 RHAQ2	gonadoliberin II
6	40	21.9	10	1 A61126	gonadoliberin - sp
7	40	21.9	10	2 A46030	gonadoliberin I -
8	40	21.9	10	2 B46030	gonadoliberin II -
9	40	21.9	30	2 F70254	hypothetical prote
10	37	20.2	10	2 A49187	gonadotropin-relea
11	37	20.2	27	2 I54247	GTP-binding regula
12	36	19.7	35	2 B85708	unknown protein en
13	35	19.1	32	2 A61624	apolipoprotein-III
14	33	18.0	36	2 A56634	neuropeptide F - A
15	32.5	17.8	21	2 C60275	fibrillin - Porph
16	32	17.5	17	4 I51887	hypothetical EWSR1
17	32	17.5	29	2 S54340	diazepam binding i
18	31.5	17.2	21	2 D60275	fibrillin - Porph
19	31	16.9	21	2 C39543	collagen alpha 3(I
20	31	16.9	25	2 PQ0683	photosystem I 16.1
21	31	16.9	26	2 E27011	MHC class II-associ
22	30	16.4	20	2 S00315	photosystem I chai
23	30	16.4	30	2 S72626	small-cell-variant
24	30	16.4	35	2 B41161	29K antigen PEB2 -
25	30	16.4	35	2 S55309	glutathione transf
26	29.5	16.1	30	2 S11613	ribosomal protein
27	29.5	16.1	33	2 PH1736	Ig heavy chain v r
28	29.5	16.1	34	2 PH1746	Ig heavy chain v r
29	29.5	16.1	34	2 PH1747	Ig heavy chain v r

30 29.5 16.1 34 2 PH1749
31 29.5 16.1 35 2 S07131
32 29.5 16.1 36 2 PH1751
33 29.5 16.1 36 2 PH1752
34 29.5 16.1 36 2 PH1753
35 29 15.8 18 2 S09723
36 29 15.8 20 2 A39543
37 29 15.8 20 2 S06149
38 29 15.8 26 2 S78761
39 29 15.8 30 2 A49955
40 29 15.8 32 2 A27074
41 29 15.8 36 2 A69287
42 28.5 15.6 22 2 PQ0070
43 28 15.3 20 2 S06150
44 28 15.3 24 2 S40139
45 28 15.3 27 2 S29211

ALIGNMENTS

RESULT 1

RHPGG

gonadoliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Note: the synthetic and natural hormones have the same physicochemical and biolog

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Comment: this hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match

Best Local Similarity 31.7%; Score 58; DB 1; Length 10;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10

|||||

RESULT 2

RHSHG

gonadoliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing Ea

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

A:Residues: 1-10 <BUR>
A:Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 31.7%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.074; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSYGLRPG 10

RESULT 3
RHAQ1
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regl. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 29.5%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.27; 0; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSYGLQPG 10

RESULT 4
A21114
gonadoliberin - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match 24.6%; Score 45; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 4.6; 2; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSYGLWPG 10

RESULT 5
RHAQ2
gonadoliberin II - American alligator
N:Alternate names: gonadotropin-releasing hormone II

C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: B60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regl. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: B60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 21.9%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 23; 2; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 2;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSHGWPG 10

RESULT 6
A61126
gonadoliberin - spotted ratfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Hydrolagus colliei (spotted ratfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C:Accession: A61126
R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.-G.; Gen. Comp. Endocrinol. 82, 152-161, 1991
A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a holbrooki
A:Reference number: A61126; MUID:91340067
A:Accession: A61126
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Experimental source: brain
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 21.9%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 23; 2; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 2;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSHGWPG 10

RESULT 7
A46030
gonadoliberin I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 21.9%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 23;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
||||:| |
Db 2 HWSHGWLP 10

RESULT 8

B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Keywords: hormone; pyroglutamic acid
F:I/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 21.9%; Score 40; DB 2; Length 10;

Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
||||:| |
Db 2 HWSHGWYP 10

RESULT 9

F70254
hypothetical protein BBK29 - Lyme disease spirochete plasmid K/lp36
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: F70254
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: F70254
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-30 <KLE>
A:Cross-references: GB:AE000788; NID:g2690123; PIDN:AAC66163.1; PID:g2690155; TIGR:BBK29
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 21.9%; Score 40; DB 2; Length 30;

Best Local Similarity 33.3%; Pred. No. 76;
Matches 8; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 10 GSSGSLDEKKIAKMEKASSVFNV 33
|: | |::| | | |:
Db 2 GNPSSGDKKMKIRKLSYYKI 25

RESULT 10

A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993

A:Title: Primary structure and biological activity of a third gonadotropin-releasing
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 20.2%; Score 37; DB 2; Length 10;

Best Local Similarity 55.6%; Pred. No. 59;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
||||:| |
Db 2 HWSHDKPG 10

RESULT 11

I54247
GTP-binding regulatory protein Gt alpha-2 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000
C:Accession: I54247
R:Hirano, T.
Hokkaido Igaku Zasshi 68, 885-893, 1993
A:Title: [Analysis of cell specific transcription of the human cone transducin alpha
A:Reference number: I54247; MUID:94156312
A:Accession: I54247
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-27 <RES>
A:Cross-references: GB:S70567; NID:g545549; PIDN:AAD14063.1; PID:g4261763
C:Superfamily: GTP-binding regulatory protein Gs alpha chain

Query Match 20.2%; Score 37; DB 2; Length 27;

Best Local Similarity 46.7%; Pred. No. 1.8e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 12 SGPSLDEKKIAKMEK 26
|| |::| | |
Db 3 SGASAEKELAKRSK 17

RESULT 12

B85708
unknown protein encoded by prophage CP-9330 [imported] - Escherichia coli (strain C15
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85708
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apocda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85708
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-35 <STO>
A:Cross-references: GB:AE005174; NID:gl2514999; PIDN:AA56126.1; GSPDB:GN00145; UMG:P
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z2052

Query Match 19.7%; Score 36; DB 2; Length 35;

Best Local Similarity 38.1%; Pred. No. 3.2e+02;
Matches 8; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 4 SYGLRPGSGPSLDEKKIAKM 24
|| |::| | |:
Db 15 SYPLKAGYSKQDQVDYKNVINM 35

RESULT 13

A61624
apolipophorin-III - southwestern corn borer (fragment)
C:Species: Diatraea grandiosella (southwestern corn borer)
C:Date: 21-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 03-Aug-1995
C:Accession: A61624

R:Burks, C.S.; Shelby, K.S.; Chippendale, G.M.
Insect Biochem. Mol. Biol. 22, 905-915, 1992

A:Title: Characteristics of apolipophorin-III of the southwestern corn borer, Diatraea g

A:Reference number: A61624

A:Accession: A61624

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-32 <BUR>

C:Keywords: hemolymph

Query Match 19.1%; Score 35; DB 2; Length 32;
Best Local Similarity 38.9%; Pred. NO. 4e+02;

Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 9 PGSSGPSLDEKIKAKMEK 26
|::| |::|::|

Db 3 PSTTPQDQEKKAETQK 20

RESULT 14

A56634

neuropeptide F - Artiopesthia triangulata

C:Species: Artiopesthia triangulata

C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000

C:Accession: A56634

R:Curry, W.J.; Shaw, C.; Johnston, C.F.; Thim, L.; Buchanan, K.D.

Comp. Biochem. Physiol. C 101, 269-274, 1992

A:Title: Neuropeptide F: primary structure from the tubellarian, Artiopesthia triangulat

A:Reference number: A56634; MUID:92362227

A:Accession: A56634

A:Molecule type: protein

A:Residues: 1-36 <CUR>

A:Note: sequence extracted from NCBI backbone (NCBIP:110973)

C:Keywords: amidated carboxyl end; neuropeptide

F:36/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 18.0%; Score 33; DB 2; Length 36;

Best Local Similarity 60.0%; Pred. NO. 8.6e+02;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 LRPGSSGPSLDEKKI 21
|::| |::|::|

Db 5 LRPRSSFSEDEYQI 19

RESULT 15

C60275

fibrillin - Porphyromonas gingivalis (strain FAY 19M-1) (fragment)

C:Species: Porphyromonas gingivalis

C:Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 11-Jan-2000

C:Accession: C60275

R:Lee, J.Y.; Sojar, H.T.; Bedi, G.S.; Genco, R.J.

Infect. Immun. 59, 383-389, 1991

A:Title: Porphyromonas (Bacteroides) gingivalis fibrillin: size, amino-terminal sequenc

A:Reference number: A60275; MUID:91099990

A:Accession: C60275

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-21 <LEE>

C:Superfamily: Porphyromonas fibrillin

Query Match 17.8%; Score 32.5; DB 2; Length 21;

Best Local Similarity 46.7%; Pred. NO. 5.6e+02;

Matches 7; Conservative 4; Mismatches 1; Indels 3; Gaps 1;

QY 10 GSSGPSLDEKIKAKM 24
|::| |::|::|

Db 3 GNAG---DESKVAKL 14

Search completed: October 10, 2002, 22:47:57
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:34:13 ; Search time 13 Seconds
(without alignments)
107.223 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHWSYGLRPGSGPSLDEKKIAKMEKASSVENVVNS 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 2415

Minimum DB seq length: 0

Maximum DB seq length: 36

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	54	29.5	10	GONL_ALLMI	P37041 alligator m
2	47	25.7	10	GONL_CLUPA	P81749 clupea pall
3	45	24.6	10	GON3_ONCKE	P20367 oncorhynch
4	40	21.9	10	GON2_CHICK	P37043 gallus gall
5	40	21.9	10	GONL_SQUAC	P27429 squalus aca
6	37	20.2	10	GON3_PETMA	P30948 petromyzon
7	35	19.1	32	APL3_DIAGR	P81471 diatraea gr
8	34	18.6	32	RS19_YEREN	Q56847 yersinia en
9	33.5	18.3	28	PF71_HCMVT	P24429 human cytom
10	33	18.0	36	NPF_ARTTR	P41334 artiposthi
11	32	17.5	10	GON1_CHEPR	P80677 chelyosoma
12	30	16.4	20	PSAF_PEA	P20119 pisum sativ
13	29.5	16.1	35	FAS_CAPHI	P08757 capra hircu
14	29	15.8	32	H2AZ_ONCMY	P22647 oncorhynch
15	29	15.8	36	Y297_ARCFU	O29945 archaeglob
16	28	15.3	10	GON2_CHEPR	P80678 chelyosoma
17	28	15.3	20	PSAF_MAIZE	P13193 zea mays (m
18	28	15.3	36	NEUY_GADMO	P80167 gadus morhu
19	28	15.3	36	NEUY_ONCMY	P29071 oncorhynch
20	27	14.8	20	CRP_MUSCA	P19094 mustelus ca
21	27	14.8	27	GRP_CANFA	P08989 canis fami
22	27	14.8	27	PSAF_SYNVU	P23078 synchococc
23	27	14.8	27	Y803_ARCFU	O29455 archaeglob
24	27	14.8	28	V103_VACCP	Q00334 vaccinia vi
25	27	14.8	33	ACT_DICVI	Q24733 dictyocaulu
26	26.5	14.5	27	TXA3_ANESU	P01535 anemonia su
27	26	14.2	20	TENB_ACTTE	P30834 actinia ten
28	26	14.2	22	ODPX_BOVIN	P22439 bos taurus
29	26	14.2	26	PSAF_CUCSA	P42048 cucumis sat
30	26	14.2	29	PSAF_SYNF6	P31083 synchococc
31	26	14.2	29	TLP_ACTDE	P81370 actinidia d
32	26	14.2	35	GP58_BPSP1	O48412 bacterioph
33	26	14.2	36	NEUY_RANRI	P29949 rana ridibu

34	26	14.2	36	1	PMY_PETMA	P80024 petromyzon
35	25.5	13.9	33	1	YLCH_BP82	Q37869 bacterioph
36	25	13.7	10	1	GRP_RANRI	P23260 rana ridibu
37	25	13.7	15	1	UC06_MAIZE	P80612 zea mays (m
38	25	13.7	20	1	JHBP_BOMMO	P81627 bombyx mori
39	25	13.7	21	1	RL21_HALCU	P05974 halobacteri
40	25	13.7	23	1	GRP_ONCMY	Q9ps30 oncorhynch
41	25	13.7	24	1	SODC_RANCA	P23417 rana catesb
42	25	13.7	25	1	GRP_SCYCA	P09472 scyllorhinu
43	25	13.7	25	1	H1L_WHEAT	P15871 triticum ae
44	25	13.7	27	1	GRP_CHICK	P01295 gallus gall
45	25	13.7	27	1	GRP_PTIG	P01294 sus scrofa

ALIGNMENTS

RESULT 1

GONL_ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991)
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR PIR; A60066; RHAQ1.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 29.5%; Score 54; DB 1; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.043; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 1;

QY 2 HWSYGLRPG 10

DB 2 HWSYGLQPG 10

RESULT 2

GONL_CLUPA STANDARD; PRT; 10 AA.
AC P81749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH)
DE (Luliberin I).
GN GNRH1.
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;

RA Chang, J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring."
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR PIR: B60066; RHAQ2.
 DR PIR: A61126; A61126.
 DR PIR: B46030; A61030.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION
 SQ SEQUENCE 10 AA: 1254 MW; 284B2E437871F5A3 CRC64;
 Query Match 21.9%; Score 40; DB 1; Length 10;
 Best Local Similarity 66.7%; Pred. No. 4.9;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 DB 2 HWSHGWLPG 10
 RESULT 5
 GONL_SQUAC STANDARD; PRT; 10 AA.
 ID GONL_SQUAC
 AC P27429;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
 DE (Luliberin).
 OS Squalus acanthias (Spiny dogfish).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 CC Elasmobranchii; Squalae; Squaloidei; Squalidae; Squalus.
 OX NCBI_TaxID=7797;
 [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=92335300; PubMed=1631133;
 RA Lovejoy D.A., Fischer W.H., Namvongchon S., Craig A.G.,
 RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
 RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
 RT dogfish brain provides insight into GNRH evolution."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION
 SQ SEQUENCE 10 AA: 1204 MW; 284B32337871F5A3 CRC64;
 Query Match 21.9%; Score 40; DB 1; Length 10;
 Best Local Similarity 66.7%; Pred. No. 4.9;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 DB 2 HWSHGWLPG 10
 RESULT 6
 GON3_PETMA STANDARD; PRT; 10 AA.
 ID GON3_PETMA
 AC P30948;

DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
 DE (Luliberin III).
 OS Petromyzon marinus (Sea lamprey).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 CC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=93178316; PubMed=8440174;
 RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
 RT "Primary structure and biological activity of a third gonadotropin-
 RT releasing hormone from lamprey brain."
 RL Endocrinology 132:1125-1131(1993).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION
 SQ SEQUENCE 10 AA: 1277 MW; 284B36237A1F5A3 CRC64;
 Query Match 20.2%; Score 37; DB 1; Length 10;
 Best Local Similarity 55.6%; Pred. No. 13;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 DB 2 HWSHDWKEPG 10
 RESULT 7
 APL3_DIAGR STANDARD; PRT; 32 AA.
 ID APL3_DIAGR
 AC P81471;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Apolipophorin-III (APOLP-III) (Fragment).
 OS Diatraea grandiosella (Southwestern corn borer).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 CC Pyraloidea; Pyralidae; Crambinae; Diatraea.
 OX NCBI_TaxID=61289;
 [1]
 RP SEQUENCE.
 RC TISSUE=Larval plasma;
 RA Burks C.S., Shelby K.S., Chippendale G.M.;
 RT "Characteristics of apolipophorin-III of the southwestern corn borer,
 RT Diatraea grandiosella".
 RL Insect Biochem. Mol. Biol. 22:905-915(1992).
 CC -!- FUNCTION: INCREASE OF LIPID CARRYING CAPACITY OF LIPOPHORIN BY
 CC COVERING THE EXPANDING HYDROPHOBIC SURFACE RESULTING FROM DIACYL-
 CC GLYCEROL UPTAKE.
 CC -!- SUBUNIT: EQUILIBRIUM BETWEEN A SOLUBLE MONOMER AND A BOUND
 CC LIPOPROTEIN FORM. APOLIPOPHORIN-III ASSOCIATES WITH LIPOPHORIN
 CC DURING LIPID LOADING UNTIL EACH PARTICLE CONTAINS 9 OR 14
 CC MOLECULES OF APOLIPOPHORIN-III.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR; SECRETED (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: HEMOLYMPH.
 CC -!- SIMILARITY: BELONGS TO THE INSECT APOLIPOPHORIN-III FAMILY.
 KW Lipid transport; Hemolymph.
 FT NON_TER 32 32
 SQ SEQUENCE 32 AA: 3584 MW; A6C2D6D675C0B22 CRC64;
 Query Match 19.1%; Score 35; DB 1; Length 32;

Best Local Similarity 38.9%; Pred. No. 1e+02; Mismatches 4; Indels 7; Gaps 0;

OC 9 PGSSGSPSLDEKTIKMEK 26
 DB 3 PSTTPPDQXKKAEEFK 20

RESULT 8
 RS19.YEREN
 ID RS19.YEREN STANDARD; PRT; 32 AA.
 AC Q56847;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 30S ribosomal protein S19 (Fragment).
 GN RPSS.
 OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 OC NCBI_TaxID=630;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=6471/76 / SEROTYPE O:3;
 RX MEDLINE=96382736; PubMed=8790600;
 RA Mertz A.K.H., Daser A., Skurnik M., Wiesmuller K., Braun J., Appel H.,
 RA Batsford S., Wu P., Distler A., Sieper J.;
 RT "The evolutionarily conserved ribosomal protein L23 and the cationic
 RT urease beta-subunit of Yersinia enterocolitica O:3 belong to the
 RT immunodominant antigens in Yersinia-triggered reactive arthritis;
 RT implications for autoimmunity."
 RL Mol. Med. 1:44-55(1994).
 CC -!- FUNCTION: PROTEIN S19 FORMS A COMPLEX WITH S13 THAT BINDS STRONGLY
 CC TO THE 16S RIBOSOMAL RNA (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U11251; AAC43514.1; -;
 CC InterPro; IPR002222; Ribosomal_S19.
 CC Pfam; PF00203; Ribosomal_S19; 1.
 CC ProDom; PD001012; Ribosomal_S19; 1.
 CC PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
 CC Ribosomal protein; rRNA-binding.
 CC NON_TER 32 32
 CC SEQUENCE 32 AA; 3659 MW; E534F701330F0338 CRC64;

Query Match 18.6%; Score 34; DB 1; Length 32;
 Best Local Similarity 46.7%; Pred. No. 1.4e+02;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OC 13 GPSLDEKTIKMEKA 27
 DB 8 GPFIDLHLKKVEKA 22

RESULT 9
 PP71.HCMVT
 ID PP71.HCMVT STANDARD; PRT; 28 AA.
 AC P24429;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DE 71 kDa outer matrix phosphoprotein (PP71) (Fragment).
 GN UL82.
 OS Human
 OC Herpesviridae;
 OC Virus.

Betaherpesvirinae; Cytomegalovirus.
 OC NCBI_TaxID=10363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220654; PubMed=1850902;
 RA Pande H., Campo K., Tanamachi B., Zaia J.A.;
 RT "Human cytomegalovirus strain Towne pp65 gene: nucleotide sequence
 RT and expression in Escherichia coli."
 RL Virology 182:220-228(1991).
 CC -!- SIMILARITY: BELONGS TO THE UL82 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M67443; AAA45995.1; -;
 CC Phosphorylation; Matrix protein.
 CC NON_TER 28 28
 CC SEQUENCE 28 AA; 2554 MW; 3C96A69C384425BD CRC64;

Query Match 18.3%; Score 33.5; DB 1; Length 28;
 Best Local Similarity 45.0%; Pred. No. 1.5e+02;
 Matches 9; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

OC 9 PGSSGSPSLDEKTIKMEKAS 28
 DB 8 PG-EGPSSSEAAAISEAAS 26

RESULT 10
 NPF.ARTTR
 ID NPF.ARTTR STANDARD; PRT; 36 AA.
 AC P41334;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neuropeptide F (NPF).
 OS Artiopesthia triangulata.
 OC Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
 OC Rhabditophora; Seriata; Tricladida; Terricola; Geoplanidae;
 OC Arthurdendyus.
 OC NCBI_TaxID=132421;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=92362227; PubMed=13541101;
 RA Curry W.J., Shaw C., Johnston C.F., Thim L., Buchanan K.D.;
 RT "Neuropeptide F: primary structure from the tubellarian, Artiopesthia
 RT triangulata."
 RL Comp. Biochem. Physiol. 101C:269-274(1992).
 CC -!- FUNCTION: MAY PERFORM AN IMPORTANT NEUROTRANSMITTER FUNCTION AND
 CC MAY REGULATE MUSCULAR ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: CENTRAL AND PERIPHERAL NERVOUS SYSTEM, AND
 CC MUSCULAR PHARYNX.
 CC -!- SIMILARITY: BELONGS TO THE NPY / PYY / PYY FAMILY.
 CC InterPro; IPR001955; Pancreatic_horm.
 CC Pfam; PF00159; hormone3; 1.
 CC SMART; SM00309; PAH; 1.
 CC PROSITE; PS00265; PANCREATIC_HORMONE_1; 1.
 CC PROSITE; PS0276; PANCREATIC_HORMONE_2; 1.
 CC Neuropeptide; Amidation.
 CC MOD_RES 36 36
 CC SEQUENCE 36 AA; 4433 MW; D969E4C464B7E70C CRC64;

Query Match 18.0%; Score 33; DB 1; Length 36;
 Best Local Similarity 60.0%; Pred. No. 2.3e+02;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OC 7 LRPSSGSPSLDEKTI 21

"The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*.";

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:35:44 ; Search time 29 Seconds
(without alignments)
214.752 Million cell updates/sec

Title: us-09-848-834A-16

Perfect score: 183
Sequence: 1 XHWSVGLRPGSSPSLDEKIKAKMEKASSVFNVNVS 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 18951

Minimum DB seq length: 0
Maximum DB seq length: 36

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp-unclassified.*
- 15: sp-rvivirus.*
- 16: sp-bacteriap.*
- 17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	24.6	33	Q9PT34	Q9pt34 oncorhynchu
2	45	24.6	33	Q9W7G0	Q9w7g0 oncorhynchu
3	40	21.9	30	O50832	O50832 borrelia bu
4	39.5	21.6	30	Q9R5S3	Q9r5s3 leuconostoc
5	38	20.8	20	Q9EQX8	Q9eqx8 mus musculu
6	38	20.8	32	Q9U340	Q9u340 caenorhabdi
7	37	20.2	27	Q16162	Q16162 homo sapien
8	35.5	19.4	31	Q9GKI8	Q9gki8 sus scrofa
9	33	18.0	29	Q95DD0	Q95dd0 acer buerge
10	33	18.0	34	Q69903	Q69903 human immun
11	33	18.0	34	Q69904	Q69904 human immun
12	32.5	17.8	19	Q9UQG7	Q9uqg7 homo sapien
13	32.5	17.8	20	Q9QVF6	Q9qvfv6 rattus sp.
14	32	17.5	31	Q9TWK5	Q9twk5 mytilus edu
15	32	17.5	33	Q95LC4	Q95lc4 sus scrofa
16	31	16.9	25	Q9W213	Q9w213 nicotiana s

17	31	16.9	27	7	Q31210	Q31210 mus musculu
18	31	16.9	28	15	Q902G2	Q902g2 human immun
19	31	16.9	29	15	Q9IQJ9	Q9iqj9 human immun
20	31	16.9	29	15	Q9IQI3	Q9iqi3 human immun
21	31	16.9	35	15	Q9QQY5	Q9qqy5 human immun
22	30	16.4	20	2	O34197	O34197 rickettsia
23	30	16.4	20	6	Q9TR34	Q9tr34 ovis aries
24	30	16.4	23	2	Q9R5R3	Q9r5r3 photobacter
25	30	16.4	24	4	Q96H37	Q96h37 homo sapien
26	30	16.4	25	4	Q9UC30	Q9uc30 homo sapien
27	30	16.4	25	5	Q9N5B6	Q9n5b6 caenorhabdi
28	30	16.4	28	3	Q07198	Q07198 saccharomyc
29	30	16.4	30	2	Q45966	Q45966 coxiella bu
30	30	16.4	32	11	Q35050	Q35050 rattus norv
31	30	16.4	33	4	Q9UD11	Q9ud11 homo sapien
32	30	16.4	34	4	Q14341	Q14341 homo sapien
33	30	16.4	35	2	Q9R5T9	Q9r5t9 campylobact
34	29.5	16.1	21	2	Q9X3G6	Q9x3g6 prochloroco
35	29.5	16.1	29	6	Q9N2H2	Q9n2h2 bos taurus
36	29.5	16.1	31	4	Q96FC8	Q96fc8 homo sapien
37	29	15.8	15	11	Q9QUZ3	Q9quz3 rattus sp.
38	29	15.8	18	13	Q90791	Q90791 gallus gall
39	29	15.8	20	11	Q9QUZ4	Q9quz4 rattus sp.
40	29	15.8	24	4	Q9NQY4	Q9ndy4 homo sapien
41	29	15.8	25	10	Q940X5	Q940x5 glycine max
42	29	15.8	29	15	Q9IQJ4	Q9iqj4 human immun
43	29	15.8	29	15	Q9IQJ3	Q9iqj3 human immun
44	29	15.8	29	15	Q9IQJ2	Q9iqj2 human immun
45	29	15.8	29	15	Q9IQI7	Q9iqi7 human immun

ALIGNMENTS

RESULT 1

Q9PT34	Q9PT34	PRELIMINARY;	PRT;	33 AA.
ID	Q9PT34;			
AC	Q9PT34;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)			
DE	(LULIBERIN) (FRAGMENT).			
GN	GNRH.			
OS	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;			
OC	Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.			
OX	NCBI_TaxID=8022;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99312119; PubMed=10385393;			
RA	Von Schalburg K.R.; Sherwood N.M.;			
RT	"Regulation and expression of gonadotropin-releasing hormone gene			
RT	differs in brain and gonads in rainbow trout.";			
RL	Endocrinology 140:3012-3024(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	von Schalburg K.R.; Sherwood N.M.;			
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.			
CC	- - FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY			
CC	SIMILARITY).			
CC	- - SIMILARITY: BELONGS TO THE GNRH FAMILY.			
DR	EMBL: AF110533; RAD43461.1; -.			
DR	InterPro: IPR002047; AKH.			
DR	InterPro: IPR002012; GNRH.			
DR	Pfam: PF00446; GNRH; 1.			
DR	PROSITE: PS00256; AKH; UNKNOWN_1.			
DR	PROSITE: PS00473; GNRH; 1.			
KW	Amidation; Hormone.			
FT	NON_TER 33			
SEQ	SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;			

RA Makinae K., Kobayashi T., Kobayashi T., Sinkawa H., Sakagami H.,
RT Kondo H., Tashiro F., Miyazaki J., Obata K., Tamura S., Yanagawa Y.,
RT "Structure of the Mouse Glutamate Decarboxylase 65 Gene and its
RT Promoter".
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032757; BAB20415.1; -.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 1989 MW; 28F544B0E4C3768D CRC64;

Query Match 20.8%; Score 38; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 WSYGLRPGSSGP 14
II:I II:I
Db 9 WSEGSSEDSADP 20

RESULT 6
Q9U340
ID Q9U340 PRELIMINARY; PRT; 32 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2001 (TREMBLrel. 16, Last annotation update)
DE W06G6.9 PROTEIN.
GN W06G6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z83129; CAB6325.1; -.
SQ SEQUENCE 32 AA; 3775 MW; E49D031288C6806A CRC64;

Query Match 20.8%; Score 38; DB 5; Length 32;
Best Local Similarity 41.7%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSG 13
II: II: II:
Db 18 HWQWAAKPGEWG 29

RESULT 7
Q16162
ID Q16162 PRELIMINARY; PRT; 27 AA.
AC Q16162;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CONE TRANSDUCIN ALPHA SUBUNIT PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94156312; PubMed=8112713;
RA Hirano T.;
RT "[Analysis of cell specific transcription of the human cone transducin
RT alpha subunit gene]."
RL Hokkaido Igaku Zasshi 68:885-893(1993).

DR EMBL; S70567; AAD14063.1; -.
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2949 MW; FBE217D50F3FC584 CRC64;

Query Match 20.2%; Score 37; DB 4; Length 27;
Best Local Similarity 46.7%; Pred. No. 2.7e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 12 SGPSLDEKKIAKMEK 26
II:II:II:II:
Db 3 SCASAEDEKELAKRSK 17

RESULT 8
Q9GKI8
ID Q9GKI8 PRELIMINARY; PRT; 31 AA.
AC Q9GKI8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ZINC FINGER DNA BINDING PROTEIN 89 (FRAGMENT).
GN ZBP-89.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Van Poucke M., Yerle M., Tuggle C., Plumi F., Van Zeveren A.,
RA Peelman L.J.;
RT "Integration of porcine chromosome 13 maps."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229222; AAG41135.1; -.
FT NON_TER 1 1
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 3442 MW; FF3CC41F9A054E0E CRC64;

Query Match 19.4%; Score 35.5; DB 6; Length 31;
Best Local Similarity 50.0%; Pred. No. 5.2e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 11 SSGP---SLDEKKIAKMEKASS 29
I: I: II:II:II:II:
Db 4 STSPKDNSLPKKRKQTEKKSS 25

RESULT 9
Q95DD0
ID Q95DD0 PRELIMINARY; PRT; 29 AA.
AC Q95DD0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ATPASE B SUBUNIT (FRAGMENT).
GN ATPB.
OS Acer buergerianum (trident maple).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Sapindales; Sapindaceae; Acer.
OX NCBI_TaxID=57649;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheng Y.P., Chiang T.Y., Chou C.H., Chiang Y.C.;
RT "atpB-zbcl intergenic spacer from Acer buergerianum Miq. var.
RT formosanum (Hayata) Sasaki."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ245942; CAC67772.1; -.
KW Chloroplast. 29 29
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3111 MW; F1431AADD86CFCD3 CRC64;

Query Match 18.0%; Score 33; DB 8; Length 29;

```
Best Local Similarity 31.6%; Pred. No. 1.1e+03;
Matches 6; Conservative 8; Mismatches 3; Indels 2; Gaps 1;

QY 8 RPSGSPSLDEKKTAKMEK 26
Db 9 RGVSV--TIEKSVGRIDQ 25

RESULT 10
Q69903 PRELIMINARY; PRT; 34 AA.
AC Q69903;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE GP41 (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94211861; PubMed=7512731;
RA Conley A.J., Kessler J.A., Boots L.J., Tung J.S., Arnold B.A.,
RA Keller P.M., Shaw A.R., Emini E.A.;
RT "Neutralization of divergent human immunodeficiency virus type 1
RT variants and primary isolates by IAM-41-2F5, an anti-gp41 human
RT monoclonal antibody.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:3348-3352(1994).
DR EMBL; U06733; AAA19146.1; -.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 34 34
SQ SEQUENCE 34 AA; 4236 MW; E609273F6164BD80 CRC64;

Query Match 18.0%; Score 33; DB 15; Length 34;
Best Local Similarity 31.2%; Pred. No. 1.3e+03;
Matches 5; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 17 DEKKIAKMEKASSVFN 32
Db 17 NEKELLELDKWSLWN 32

RESULT 11
Q69904 PRELIMINARY; PRT; 34 AA.
AC Q69904;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE GP41 (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrov. viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94211861; PubMed=7512731;
RA Conley A.J., Kessler J.A., Boots L.J., Tung J.S., Arnold B.A.,
RA Keller P.M., Shaw A.R., Emini E.A.;
RT "Neutralization of divergent human immunodeficiency virus type 1
RT variants and primary isolates by IAM-41-2F5, an anti-gp41 human
RT monoclonal antibody.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:3348-3352(1994).
DR EMBL; U06734; AAA19147.1; -.
DR InterPro; IPR000328; Env_GP41.
DR Pfam; PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1 1
FT NON_TER 34 34
SQ SEQUENCE 34 AA; 4236 MW; E609273F6164BD80 CRC64;
```

```
Best Local Similarity 31.2%; Pred. No. 1.3e+03;
Matches 5; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 17 DEKKIAKMEKASSVFN 32
Db 17 NEKELLELDKWSLWN 32

RESULT 12
Q9UQG7 PRELIMINARY; PRT; 19 AA.
AC Q9UQG7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE ENDOPLASMIC RETICULUM-GOLGI INTERMEDIATE COMPARTMENT PROTEIN ERGIC-53
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99192442; PubMed=10090935;
RA Nichols W.C., Terry V.H., Wheatley M.A., Yang A., Zivelin A.,
RA Ciavarella N., Stefanie C., Matsushita T., Saito H., de Bosch N.B.,
RA Ruiz-Saez A., Torres A., Thompson A.R., Feinstein D.I., White G.C.,
RA Negrier C., Vinciguerra C., Aktan M., Kaufman R.J., Ginsburg D.,
RA Seligsohn U.;
RT "ERGIC-53 gene structure and mutation analysis in 19 combined factors
RT V and VIII deficiency families.";
RL Blood 93:2261-2266(1999).
DR EMBL; AF081877; AAD32485.1; -.
DR EMBL; AF081876; AAD32485.1; JOINED.
FT NON_TER 1 1
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2166 MW; 7BDB4711437A50F5 CRC64;

Query Match 17.8%; Score 32.5; DB 4; Length 19;
Best Local Similarity 44.4%; Pred. No. 8.2e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 9 PGSSGSPSLDEKKTAKMEK 26
Db 1 PGKEPPTPD-KEISEKEK 17

RESULT 13
Q9QVF6 PRELIMINARY; PRT; 20 AA.
AC Q9QVF6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CARBONIC ANHYDRASE IV, CA IV.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=92147688; PubMed=1737787;
RA Waheed A., Zhu X.L., Sly W.S.;
RT "Membrane-associated carbonic anhydrase from rat lung. Purification,
RT characterization, tissue distribution, and comparison with carbonic
RT anhydrase IVs of other mammals.";
RL J. Biol. Chem. 267:3308-3311(1992).
SQ SEQUENCE 20 AA; 2275 MW; 6CB2AC1CD4DEB65 CRC64;

Query Match 17.8%; Score 32.5; DB 11; Length 20;
```

Best Local Similarity 44.4%; Pred. NO. 8.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 5; Gaps 2;

QY 2 HWSYGLR---PGS--SGP 14
Db 3 HWXYEIQAKEPNSXXSGP 20

RESULT 14

Q9TWK5
ID Q9TWK5 PRELIMINARY; PRT; 31 AA.

AC Q9TWK5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PROXIMAL COLLAGEN, COL-P.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=95230211; PubMed=7714453;
RA Qin X., Waite J.H.;
RT "Exotic collagen gradients in the byssus of the mussel Mytilus
RT edulis.";
RL J. Exp. Biol. 198;633-644(1995).
DR InterPro; IPR000087; Collagen
SQ SEQUENCE 31 AA; 2648 MW; B1F7708959101A73 CRC64;

Query Match 17.5%; Score 32; DB 5; Length 31;
Best Local Similarity 71.4%; Pred. NO. 1.7e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 PGSSGPS 15
Db 15 PGTCPT 21

RESULT 15

Q95LC4
ID Q95LC4 PRELIMINARY; PRT; 33 AA.

AC Q95LC4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BETA-1 ADRENERGIC RECEPTOR (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21540585; PubMed=11683720;
RA Lee J.H., Zhang W., Moran C.;
RT "Comparative porcine gene mapping relative to human chromosomes 9, 10,
RT 20, and 22.";
RL Anim. Genet. 32:313-315(2001).
DR EMBL; AF345500; AAK57992.1; -.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3694 MW; 7DBEBDC3E268E150 CRC64;

Query Match 17.5%; Score 32; DB 6; Length 33;
Best Local Similarity 80.0%; Pred. NO. 1.8e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYG 6
Db 11 HWDYG 15

Search completed: October 10, 2002, 22:47:20
Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:33:08 ; Search time 27 Seconds
(without alignments)
148.098 Million cell updates/sec

Title: US-09-848-834A-16

Perfect score: 183

Sequence: 1 XHWSYGLRPGSGPSLDEKIKAKMEKASSVFNVNS 36

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 307298

Minimum DB seq length: 0

Maximum DB seq length: 36

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89.5	48.9	33	22	AA63663 Peptide comprising
2	89.5	48.9	33	22	AA63516 A peptide which ma
3	89	48.6	19	22	AAW98951 Vaccine related MH
4	89	48.6	21	10	AAAP91504 Sequence of modifi
5	89	48.6	21	16	AAAR82586 Plasmodium falcipa
6	89	48.6	21	16	AAAR78920 Malaria circumspor
7	89	48.6	21	16	AAAR75955 p. falciparum CS p
8	89	48.6	21	16	AAAR70912 Malaria circumspor
9	89	48.6	21	17	AAW05612 Circumsporozoite h
10	89	48.6	21	18	AAW35440 T-cell stimulatory
11	89	48.6	21	20	AAAY23252 Peptide derived fr

12	89	48.6	21	21	AAAY80071 Pathogen derived T
13	89	48.6	21	21	AAAY54553 T helper cell (Th)
14	89	48.6	21	21	AAAY58777 Unidentified pepti
15	89	48.6	21	22	AAAB99706 Plasmodium falcipa
16	89	48.6	21	22	AAAG62428 Plasmodium falcipa
17	89	48.6	21	22	AAAG84517 Plasmodium falcipa
18	89	48.6	21	22	AAAG82669 Plasmodium falcipa
19	89	48.6	21	22	AAAG89366 Plasmodium falcipa
20	89	48.6	21	22	AAAB84447 Sequence of T help
21	89	48.6	21	22	AAAB98457 Plasmodium falcipa
22	89	48.6	22	16	AAAR82077 Malaria CST3 prote
23	89	48.6	33	15	AAAR62715 LHRH-containing im
24	89	48.6	33	16	AAAR83570 IgE CH4 region con
25	85	46.4	18	21	AAAY92599 CD4+ T cell epitop
26	82	44.8	21	15	AAAR65375 Helper T cell epit
27	81	44.3	21	21	AAAY70283 Plasmodium falcipa
28	79	43.2	17	16	AAAR78919 Malaria circumspor
29	79	43.2	17	16	AAAR70911 Malaria circumspor
30	79	43.2	17	21	AAAY90332 HLA class II bindi
31	79	43.2	17	22	AAAM98950 Vaccine related MH
32	79	43.2	19	21	AAAY90333 HLA class II bindi
33	75	41.0	16	22	AAAB46168 Plasmodium sp mala
34	75	41.0	16	22	AAAB49067 Malaria CS T3 T-ce
35	75	41.0	17	22	AAAM98933 Vaccine related MH
36	75	41.0	17	22	AAAM98935 Vaccine related MH
37	75	41.0	17	22	AAAM98936 Vaccine related MH
38	75	41.0	17	22	AAAM98938 Vaccine related MH
39	75	41.0	17	22	AAAM98957 Vaccine related MH
40	75	41.0	17	22	AAAM98961 Vaccine related MH
41	75	41.0	17	22	AAAM98964 Vaccine related MH
42	74.5	40.7	20	22	AAJ04118 P falciparum CS pr
43	74	40.4	17	22	AAAM98956 Vaccine related MH
44	74	40.4	17	22	AAAM98962 Vaccine related MH
45	74	40.4	17	22	AAAM98963 Vaccine related MH

ALIGNMENTS

RESULT 1

AA63663
ID AA63663 standard; peptide; 33 AA.

AC AA63663;

DT 29-OCT-2001 (first entry)

XX Peptide comprising conjugation sites for a pseudopeptide.

DE Pseudopeptide; immunomodulation; adjuvant; antigen presenting cell;

KW macrophage; dendritic cell; vaccine; autoimmune disease.

XX Synthetic.

OS WO200146127-A1.

PN 28-JUN-2001.

XX 22-DEC-1999; 99WO-IB02038.

XX 22-DEC-1999; 99WO-IB02038.

XX (OMPH-) OM-PHARMA.

XX Bauer J, Martin OR, Rodriguez S;

PI WPI; 2001-502469/55.

XX New amphiphilic acylated pseudopeptides having a functionalized

PT auxiliary spacer, useful as immunomodulators, e.g. as adjuvants in

PT vaccines

XX Example 3; Page 61; 166pp; French.

PS

CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
 CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful, in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
 CC allergy, graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
 CC or fungal infections; or (ii) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I), e.g. melanoma
 CC and human immunodeficiency virus infection. AAM98898 to AAM99592
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.

XX Sequence 19 AA;

Query Match 48.6%; Score 89; DB 22; Length 19;
 Best Local Similarity 100.0%; Pred. No. 3.9e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIARMEKASSVFNVNS 36
 Db 1 EKKIARMEKASSVFNVNS 19
 |||||

RESULT 4

AAP91504
 ID AAP91504 standard; peptide; 21 AA.

XX
 AC AAP91504;

DT 13-MAR-1992 (first entry)

DE Sequence of modified Plasmodium circumsporozoite T-cell epitope CS.T3.
 KW Circumsporozoite peptide; T-cell epitope; immunogenic composition;
 KW vaccine.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 1..2

FT /note= "May be H-Asp-Ile, H-Ile, or H-"

FT Misc-difference 19..21

FT /note= "May be Val-Asn-Ser-OH, Val-Asn-OH,
 Val-OH or -OH "

FT

XX EP343460-A.

XX PD 29-NOV-1989.

XX PF 12-MAY-1989; 89EP-0108618.

XX PR 24-MAY-1988; 88GB-0012214.

XX PA (HOFF) HOFFMANN-LA ROCHE AG.

XX PI Sinigaglia F;

XX WPI; 1989-349561/48.

XX Modified Plasmodium CS peptide - used as a universally recognised
 PT T-cell epitope in vaccines to elicit an immune response against
 PT pathogenic agents
 XX Claim 1; Page 16; 23pp; English.

XX Peptide CS.T3 (residues 1-21) corresponds to residues 378-398 of the
 CC CS protein from P. falciparum but contains 2 Ala residues in place
 CC of the native protein's Cys residues at positions 384 and 389. Also

CC claimed is AAP91504 (or modified forms, see FT) associated with an
 CC antigenic structure representing a B-cell epitope, pref. a multiple
 CC antigenic peptide, esp. multimers of the repeat sequences NAMP
 CC present in P. falciparum CS protein.

XX Sequence 21 AA;

Query Match 48.6%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIARMEKASSVFNVNS 36
 Db 3 EKKIARMEKASSVFNVNS 21
 |||||

RESULT 5

AAR82586

ID AAR82586 standard; peptide; 21 AA.

XX
 AC AAR82586;

DT 13-JUN-1996 (first entry)

DE Plasmodium falciparum circumsporozoite helper T cell epitope, PF.

KW IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 KW vaccine; allergy; antibody; constant heavy chain.

XX Plasmodium falciparum.

XX WO9526365-A1.

XX PD 05-OCT-1995.

XX PF 24-MAR-1995; 95WO-US03741.

XX PR 25-OCT-1994; 94US-0328912.

XX PR 28-MAR-1994; 94US-0218461.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX WPI; 1995-351297/45.

XX Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper
 PT T cell epitope - useful for eliciting antibody prodn. for allergy
 PT treatment
 XX Claim 3; Page 23; 87pp; English.

XX AAR82571-91 are helper T cell epitopes which can be used in the
 CC preparation of a peptide immunogen that is useful in vaccines for
 CC treating allergic reactions. In the immunogen an IgE CH4 peptide
 CC is attached C-terminally to a series of amino acids including a
 CC helper T cell epitope. The immunogen may also opt. contain a fatty
 CC acid or fatty acid derivative, an invasion domain or alpha-NH2. The
 CC immunogen produces high titres of antibodies to the effector site
 CC in human IgE heavy chain (the CH4 domain peptide) which inhibit mast
 CC cell activation and reduce allergen-induced IgE prodn. The immunogens
 CC may be used in either a radially branching multimeric form or a
 CC linearly arranged monomeric form.

XX Sequence 21 AA;

Query Match 48.6%; Score 89; DB 16; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIARMEKASSVFNVNS 36
 Db 3 EKKIARMEKASSVFNVNS 21
 |||||

RESULT 6

AAR78920
ID AAR78920 standard; peptide; 21 AA.

XX AC AAR78920;

XX DT 27-MAR-1996 (first entry)

XX DE Malaria circumsporozoite 378-398 helper T lymphocyte epitope.

XX KW 378-398; cytotoxic T; CTL; epitope; helper T; HTL; cell;

XX KW lymphocyte; antigens; treatment; disease prevention; tumours;

XX KW viruses; parasites; malaria circumsporozoite; hepatitis B.

XX OS Malaria circumsporozoite.

XX PN WO9522317-A1.

XX PD 24-AUG-1995.

XX PF 16-FEB-1995; 95WO-US02121.

XX PR 16-FEB-1994; 94US-0197484.

XX PA (CYTE-) CYTEL CORP.

XX PI Ceut RW, Grey H, Sette AD, Vitiello MA;

XX DR WPI; 1995-302545/39.

XX PT Compn. inducing cytotoxic T lymphocyte response to pref. viral,
bacterial, parasitic or tumour antigens - useful in the treatment
and prevention of diseases associated with the antigen e.g.
hepatitis B

XX PS Disclosure; Page 29; 109pp; English.

XX CC A compn. which induces a cytotoxic T lymphocyte (CTL) response to
an antigen (Ag) in a mammal comprises a CTL Ag response inducing
peptide and a lipid conjugated helper T cell inducing peptide (i.e.
AAR78918-R78922). The compn. induces a CTL response to bacterial,
viral or tumour Ags, and is therefore useful in the treatment and
prevention of diseases associated with the Ag, e.g. hepatitis B.

XX SQ Sequence 21 AA;

Query Match 48.6%; Score 89; DB 16; Length 21;

Best Local Similarity 100.0%; Pred. NO. 4.4e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKIKAKMERASSVFNVNS 36

DB 3 EKIKAKMERASSVFNVNS 21

RESULT 7

AAR75955
ID AAR75955 standard; Peptide; 21 AA.

XX AC AAR75955;

XX DT 06-MAR-1996 (first entry)

XX DE P. falciparum CS protein epitope residues 378-398.

XX KW MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;
cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.

XX OS Plasmodium falciparum.

XX PN WO9519783-A1.

XX

PD 27-JUL-1995.

XX PF 25-JAN-1995; 95WO-US01000.

XX PR 25-JAN-1994; 94US-0186266.

XX PA (CYTE-) CYTEL CORP.

XX PI Celis E, Grey HM, Kubo RT, Sette A;

XX DR WPI; 1995-269270/35.

XX PT Immunogenic peptide(s) that induce immune response to cancer cells
- that express a MAGE-3 protein peptide epitope used in vaccines or
adoptive immuno-therapy to induce cytotoxic T lymphocytes

XX PS Disclosure; Page 14; 44pp; English.

XX CC AAR75942 is derived from the sequence of the melanoma antigen (MAGE-3)
protein and can be used to elicit a primary cytotoxic T lymphocyte
response against cells expressing MAGE-3. Synthetic peptides AAR75945-53
can be used therapeutically to elicit CTL responses to melanoma, breast,
colon, prostate, or other cells which express proteins with this
epitope. The peptides have specific HLA-A1 binding capacity. The peptides
can be also used in vaccines, esp. combined with peptides such as
AAR75955-56, which are T-helper epitopes.

XX SQ Sequence 21 AA;

Query Match 48.6%; Score 89; DB 16; Length 21;

Best Local Similarity 100.0%; Pred. NO. 4.4e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKIKAKMERASSVFNVNS 36

DB 3 EKIKAKMERASSVFNVNS 21

RESULT 8

AAR70912
ID AAR70912 standard; Protein; 21 AA.

XX AC AAR70912;

XX DT 09-OCT-1995 (first entry)

XX DE Malaria circumsporozoite 378-398 T helper peptide.

XX KW Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;

XX KW malaria circumsporozoite 378-398; T helper peptide.

XX OS Malaria circumsporozoite.

XX PN WO9504542-A.

XX PD 16-FEB-1995.

XX PF 02-AUG-1994; 94WO-US08721.

XX PR 06-AUG-1993; 93US-0103623.

XX PA (CYTE-) CYTEL CORP.

XX PI Fikes JD, Livingston BD, Sette AD, Sidney JC;

XX DR WPI; 1995-090681/12.

XX PT Human melanoma antigen, MAGE-1, peptide(s) - useful for
stimulating immune response against melanoma.

XX PS Disclosure; Page 13; 59pp; English.

XX

CC The T helper peptides described in AAR70910-R70914 are used in
 CC conjunction with the C-terminal MAGE-1 peptides described in
 CC AAR70915 to AAR70969. Compens. containing the T helper and MAGE-1
 CC peptides can be administered, as a vaccine to patients susceptible
 CC to MAGE associated tumours, e.g. melanomas.
 XX
 SQ Sequence 21 AA; DB 16; Length 21; Mismatches 0; Indels 0; Gaps 0;
 Query Match 48.6%; Score 89; DB 16; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 18 EKKIAMEKASSVFNVNS 36
 DB 3 EKKIAMEKASSVFNVNS 21
 RESULT 9
 AAW05612
 ID AAW05612 standard; peptide; 21 AA.
 XX
 AC AAW05612;
 XX
 DT 10-DEC-1996 (first entry)
 DE Circumsporozoite helper T cell epitope.
 XX
 KW Immunoglobulin; IgE; membrane protein; human; epsilon chain; hepatitis B;
 KW membrane anchoring domain; helper T cell; surface antigen; core antigen;
 KW pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;
 KW Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;
 KW diphtheria toxin; plasmodium falciparum; circumsporozoite; E. coli Trat;
 KW schistosoma mansoni; triose phosphate isomerase; allergenic reaction;
 KW allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;
 KW antihistamine; decongestant; beta-2 agonist; immunosuppression;
 KW corticosteroid.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO9612740-A1.
 XX
 PD 02-MAY-1996.
 XX
 PF 25-OCT-1995; 95WO-US13841.
 XX
 PR 25-OCT-1994; 94US-0328519.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Walfield AM, Wang CV;
 XX
 DR WPI; 1996-230555/23.
 XX
 PT Peptide immunogen useful in treatment of allergy - comprises
 PT membrane-bound IgE epsilon-chain peptide synthesised linearly in
 PT tandem with T helper epitope peptide
 XX
 PS Claim 2; Page 19; 53pp; English.
 XX
 CC AAW05957-W05616 represent helper T cell epitopes used in the peptide
 CC immunogens of the invention. This sequence represents the Plasmodium
 CC falciparum circumsporozoite helper T cell antigen. The peptides of the
 CC invention contain one of these sequences, and a membrane-bound
 CC immunoglobulin E (IgE) fragment (see AAW05595 and AAW05596). The
 CC peptide immunogens of the invention can be used in vaccines for the
 CC immunotherapeutic treatment of allergic reactions, including allergic
 CC rhinitis, food allergies, anaphylaxis, or virally-induced asthma. The
 CC immunogens overcome the short effective period of antihistamines,
 CC decongestants, and beta-2 agonists, while preventing the broad
 CC immunosuppression of corticosteroids. The peptides do not have the
 CC potential side effects of restlessness or sedation (associated with
 CC antihistamines), associated increased morbidity in asthmatics (as seen
 CC with beta-2 agonists) and adverse hormonal activities (observed in

CC corticosteroid users).
 XX
 SQ Sequence 21 AA; DB 17; Length 21; Mismatches 0; Indels 0; Gaps 0;
 Query Match 48.6%; Score 89; DB 17; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 18 EKKIAMEKASSVFNVNS 36
 DB 3 EKKIAMEKASSVFNVNS 21
 RESULT 10
 AAW35440
 ID AAW35440 standard; peptide; 21 AA.
 XX
 AC AAW35440;
 XX
 DT 22-APR-1998 (first entry)
 XX
 DE T-cell stimulatory peptide from Plasmodium falciparum.
 XX
 KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
 KW scaffold; inhibition; metastasis; wound healing; solid phase.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO9738011-A1.
 XX
 PD 16-OCT-1997.
 XX
 PF 03-APR-1997; 97WO-DE00146.
 XX
 PR 03-APR-1996; 96DK-0000398.
 XX
 PA (PEPR-) PEPRESEARCH AS.
 XX
 PI Heegaard PMH, Jakobsen PH;
 XX
 DR WPI; 1997-512645/47.
 XX
 PT Non-dendritic peptide carrier linked to a solid phase - useful as a
 PT diagnostic agent and as a scaffold for production of chemical
 PT derivatives
 XX
 PS Claim 30; Page 199; 262pp; English.
 XX
 CC A non-dendritic peptide carrier (A) has been developed which is coupled
 CC through a linker to a solid phase, forming a complex of (A)-solid phase.
 CC Where (A) comprises 10-50 amino acids capable of forming a secondary
 CC structure in a benign buffer after liberation from the solid phase, and
 CC further the (A)-solid phase complex comprises an immunogenic substance
 CC and/or an immune mediator coupled on (A). The present sequence
 CC represents a specifically claimed T-cell stimulatory peptide from the
 CC present invention. An (A)-solid phase complex can be used as a scaffold
 CC for the production of chemical derivatives, characterised by covalently
 CC attaching molecules at attachment points. Alternatively (A) is used as
 CC a scaffold-peptide for the incorporation into an immunostimulating
 CC complex (Iscom) resulting in an (A)-Iscom complex which is used for the
 CC chemical coupling of antigenic substances in an aqueous solution by
 CC conjugation. (A) derivatised with one or more peptides having
 CC fibronectin-, laminin- or vitronectin-like binding activities can be
 CC used for the promotion of cell-attachment to plastic surfaces, in
 CC particular to inhibit tumour growth and metastasis, and for promotion
 CC of wound healing. Also a derivatised (A) can be used for the selection
 CC of specifically-binding aptamers or as a diagnostic agent. Such
 CC diagnostic-(A) molecules could be used to detect molecules derived from
 CC or indicative of pregnancy or of a disease, such as an infectious,
 CC autoimmune or cancerous disease.
 XX
 SQ Sequence 21 AA;

KW 123

rheumatoid ar

rheumatoid ar

XX OS Plasmodium falciparum.
 XX PN W09967294-A1.
 XX PD 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US14030.
 XX PR 20-JUN-1998; 98US-0100409.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY;
 XX DR WPI; 2000-160579/14.
 XX PT New antigenic peptide from the CDR2 domain of CD4, for immunization against e.g. human immune deficiency virus -
 XX PS Claim 11; Page 65; 106pp; English.
 XX CC The present sequence represents a broadly reactive promiscuous T helper cell (Th) epitope derived from Plasmodium falciparum circumsporozoite. It is conjugated to antigenic peptides derived from the CDR2-like domain of the human CD4 protein. These antigenic peptides present neutralising receptor/co-receptor effector sites of the CDR2-like domain. The peptides evoke effective antibody responses by having optimised site-specificity. The induced antibodies block human immune deficiency virus (HIV) binding and syncytia formation. They may also block CD4-Class II interactions with other cells, deliver signals to T cells (inhibiting normal CD4+-mediated immunoregulatory functions) or induce apoptosis of CD4 cells by simultaneous engagement of T cell receptors. Conjugates and peptides containing the antigenic peptides are used for active immunisation to generate antibodies against CD4 surface complexes, especially to prevent binding of HIV to CD4 and thus HIV infection, but also to treat undesirable immune responses such as transplant rejection, or autoimmune diseases (rheumatoid arthritis, systemic lupus erythematosus or psoriasis). These conjugates produce high-titre antibodies which are broadly neutralising against primary isolates from all classes of HIV-1 and of HIV-2.
 XX SQ Sequence 21 AA;
 Query Match 48.6%; Score 89; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 EKIKAKMEKASSVFNVNS 36
 Db 3 EKIKAKMEKASSVFNVNS 21
 RESULT 14
 AAY58777
 ID AAY58777 standard; Peptide; 21 AA.
 AC AAY58777;
 XX DT 25-APR-2000 (first entry)
 XX DE Unidentified peptide.
 KW Helper T cell; Th epitope; feed additive; growth promotion;
 KW somatostatin.
 XX OS Unidentified.
 XX PN W09966950-A1.
 XX PD 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US13923.

XX PR 20-JUN-1998; 98US-0100415.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY;
 XX DR WPI; 2000-160560/14.
 XX PT New somatostatin helper T-cell epitope conjugate for raising anti-somatostatin antibodies to enhance growth rate in animal by reducing growth inhibitory activity of somatostatin -
 XX PS Disclosure; Page 53; 59pp; English.
 XX CC The present sequence is that of an unidentified peptide of the invention. The invention relates to peptide compositions (see AAY58739-66) useful as immunogens for growth promotion in farm animals. The immunogenic peptides contain helper T cell epitopes which comprise multiple class II MHC motifs and have somatostatin at either the C- or N-terminus. They may also include an invasion domain which acts as a general immune stimulator. The helper T cell epitopes and the invasion domain enhance the immune response against the somatostatin self-peptide.
 XX SQ Sequence 21 AA;
 Query Match 48.6%; Score 89; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 EKIKAKMEKASSVFNVNS 36
 Db 3 EKIKAKMEKASSVFNVNS 21
 RESULT 15
 AAB99706
 ID AAB99706 standard; peptide; 21 AA.
 AC AAB99706;
 XX DT 06-SEP-2001 (first entry)
 XX DE Plasmodium falciparum CS protein fragment 378-398 SEQ ID NO:27.
 XX KW Human leukocyte antigen A2 binding peptide; HLA class I A2; CTL; cytotoxic T-cell lymphocyte; tumour associated antigen; CEA; HER2/neu; MAGE2; MAGE3; p53; vaccine; cancer; cytostatic; immunomodulator; immunotherapy; immune response.
 XX OS Plasmodium falciparum.
 XX PN W0200141741-A1.
 XX PD 14-JUN-2001.
 XX PF 13-DEC-2000; 2000WO-US34318.
 XX PR 13-DEC-1999; 99US-0170448.
 XX PR 05-APR-2000; 2000US-0543608.
 XX PR 30-MAY-2000; 2000US-0583200.
 XX PA (EPIM-) EPIMMUNE INC.
 XX PI Fikes J, Sette A, Sidney J, Southwood S, Celis E, Keogh E; Chesnut R;
 XX DR WPI; 2001-381489/40.
 XX PT Compositions for use in a vaccine for treating, e.g., breast, lung and colon cancer comprises at least one peptide that comprises an isolated epitope of a tumor-associated antigen -

XX
 PS Disclosure; Page 31; 86pp; English.
 XX
 CC The present invention describes a composition (I) comprising at least
 CC one peptide that comprises an isolated, prepared epitope consisting of
 CC a sequence selected from 25 short amino acid sequences given in AAB99680
 CC to AAB99704. Also described are: (1) a composition (II) comprising one
 CC or more peptides, and further comprising at least two epitopes selected
 CC from the 25 short amino acid sequences (as above), where each of the one
 CC or more peptides comprise less than 50 contiguous amino acids that have
 CC 100% identity with a native peptide sequence; and (2) a vaccine
 CC composition (III) comprising an epitope selected from the 25 short amino
 CC acid sequences (as above) and a pharmaceutical excipient. (I) has
 CC cytostatic and immunomodulatory activities and can be used in vaccine
 CC production and immunotherapy. The peptide epitope compositions (I)-(II)
 CC are useful for monitoring an immune response to a tumour associated
 CC antigen or when one or more peptides are combined to create a vaccine
 CC (III) that stimulates the cellular arm of the immune system. In
 CC particular, the vaccine mediates immune responses against tumours in
 CC individuals who bear an allele of the human leukocyte antigen (HLA)-A2
 CC supertype and improve the standard of care for patients being treated
 CC for breast, colon, or lung cancer. The present sequence represents a
 CC Plasmodium falciparum CS protein fragment which is given in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 21 AA;
 Query Match 48.6%; Score 89; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.4e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 EKKIARMEKASSVFNVVNS 36
 Db 3 EKKIARMEKASSVFNVVNS 21
 Search completed: October 10, 2002, 22:45:57
 Job time : 28 secs


```
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Plasmodium falciparum CS
; OTHER INFORMATION: protein at positions 378-398."
; US-08-186-266-6

Query Match 48.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIARMEKASSVFNVNS 36
Db 3 EKKIARMEKASSVFNVNS 21

RESULT 2
US-08-446-692-48
; Sequence 48, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-48

Query Match 48.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIARMEKASSVFNVNS 36
Db 3 EKKIARMEKASSVFNVNS 21

RESULT 3
US-08-488-351A-48
; Sequence 48, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-48

Query Match 48.6%; Score 89; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIARMEKASSVFNVNS 36
Db 3 EKKIARMEKASSVFNVNS 21

RESULT 4
US-09-100-409A-54
; Sequence 54, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
```


STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100.409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-54

Query Match 48.6%; Score 89; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAMEKASSVFNVNS 36
Db 3 EKKIAMEKASSVFNVNS 21

RESULT 5
PCT-US95-02121-97
Sequence 97, Application PC/TUS9502121
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
CTL IMMUNITY
NUMBER OF SEQUENCES: 133
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02121
FILING DATE: 16-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,484
FILING DATE: 16-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,682
FILING DATE: 29-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,568
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..21
OTHER INFORMATION: /note= "Malaria circumsporozoite
OTHER INFORMATION: 378-398"
PCT-US95-02121-97

Query Match 48.6%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAMEKASSVFNVNS 36
Db 3 EKKIAMEKASSVFNVNS 21

RESULT 6
PCT-US95-13841-20
Sequence 20, Application PC/TUS9513841
GENERAL INFORMATION:
APPLICANT: United Biomedical Inc; Walfield, Alan M.;
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Synthetic IgE Membrane Anchor
TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13841
FILING DATE: 25-OCT-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/328,519
FILING DATE: 25-OCT-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lin, Maria C.H.
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4117
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-13841-20

```
Query Match 48.6%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAMEKASSVFNVNS 36
Db 3 EKKIAMEKASSVFNVNS 21

RESULT 7
US-08-446-692-27
; Sequence 27, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulato
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-27

Query Match 48.6%; Score 89; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAMEKASSVFNVNS 36
Db 3 EKKIAMEKASSVFNVNS 21

RESULT 8
US-08-488-351A-27
; Sequence 27, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulato
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-27

Query Match 48.6%; Score 89; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAMEKASSVFNVNS 36
Db 3 EKKIAMEKASSVFNVNS 21

RESULT 9
PCT-US95-02121-96
; Sequence 96, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/874,491
;; FILING DATE: 27-APR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/827,682
;; FILING DATE: 29-JAN-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: -US 07/749,568
;; FILING DATE: 26-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parmelee, Steven W.
;; REGISTRATION NUMBER: 31,990
;; REFERENCE/DOCKET NUMBER: 14137-26-4PC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 467-9600
;; TELEFAX: (415) 543-5043
;; INFORMATION FOR SEQ ID NO: 96:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 17 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1..17
;; OTHER INFORMATION: /note="Malaria circumsporozoite"
;; OTHER INFORMATION: 382-398"
PCT-US95-02121-96

Query Match 43.2%; Score 79; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 K1AKMEKASSVFNVVNS 36
Db 1 K1AKMEKASSVFNVVNS 17
|||||

RESULT 10
US-08-817-933A-7
;; Sequence 7, Application US/08817933A
;; Patent No. 5945104
;; GENERAL INFORMATION:
;; APPLICANT: STANWORTH, DENIS R.
;; APPLICANT: LEWIN, IAN V.
;; TITLE OF INVENTION: PEPTIDES FOR ANTI-ALLERGY TREATMENT
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Nixon & Vanderhye
;; STREET: 1100 No. 5945104th Glebe Road, 8th Floor
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: US
;; ZIP: 22201-4714
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/817,933A
;; FILING DATE: 27-JUN-1997
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9422294.0
;; FILING DATE: 04-NOV-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mitchard, Leonard C.
;; REGISTRATION NUMBER: 29,009
;; REFERENCE/DOCKET NUMBER: 179-23
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 816-4005

;; TELEFAX: (703) 816-4100
;; TELEX: N/A
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 16 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-817-933A-7

Query Match 41.0%; Score 75; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EK1AKMEKASSVFNV 33
Db 1 EK1AKMEKASSVFNV 16
|||||

RESULT 11
US-08-465-167A-20
;; Sequence 20, Application US/08465167A
;; Patent No. 5750395
;; GENERAL INFORMATION:
;; APPLICANT: Fikes, John D.
;; APPLICANT: Livingston, Brian D.
;; APPLICANT: Sette, Alessandro D.
;; APPLICANT: Sidney, John C.
;; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
;; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
;; NUMBER OF SEQUENCES: 51
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, 8th Floor
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94111
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/465,167A
;; FILING DATE: 05-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/103,623
;; FILING DATE: 06-AUG-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parmelee, Steven W.
;; REGISTRATION NUMBER: 31,990
;; REFERENCE/DOCKET NUMBER: 14137-60-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-467-9600
;; TELEFAX: 415-576-0300
;; INFORMATION FOR SEQ ID NO: 20:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 20 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-465-167A-20

Query Match 40.2%; Score 73.5; DB 1; Length 20;
Best Local Similarity 94.7%; Pred. No. 0.00035;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 18 EK1AKMEKASSVFNVVNS 36
Db 1 EK1AKMEKASSVFNVVNS 36
|||||

Db 3 EKKIAKM-KASSVFNVNS 20

RESULT 12

PCT-US92-07218-17
; Sequence 17, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitello, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: PCT/US92/07218
; FILING DATE: 19920826
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14137-26-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-328-2400
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US92-07218-17

Query Match 40.2%; Score 73.5; DB 5; Length 20;
Best Local Similarity 94.7%; Pred. No. 0.00035;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 18 EKKIAKMEKASSVFNVNS 36
||||| |||||||

Db 3 EKKIAKM-KASSVFNVNS 20

RESULT 13

US-08-188-223-6
; Sequence 6, Application US/08188223
; Patent No. 5688506
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scibienski, Robert
; TITLE OF INVENTION: Immunogens Against Gonadotropin
; TITLE OF INVENTION: Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/188,223
; FILING DATE: 27-JAN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286
; TELEFAX: 212-354-8113
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /label= pGlu
; OTHER INFORMATION: /note= "Pyroglutamic acid (5-oxoproline)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..10
; OTHER INFORMATION: /note= "immunomimic"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 11..17
; OTHER INFORMATION: /note= "spacer"
; US-08-188-223-6

Query Match 38.8%; Score 71; DB 1; Length 17;
Best Local Similarity 92.3%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSGP 14
|||||

Db 2 HWSYGLRPGSSPP 14

RESULT 14

US-08-968-466-6
; Sequence 6, Application US/08968466
; Patent No. 6132720
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scibienski, Robert
; TITLE OF INVENTION: Immunogens Against Gonadotropin
; TITLE OF INVENTION: Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,466
FILING DATE: 27-JAN-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: YES

FEATURE:

NAME/KEY: Modified-site
LOCATION: 1

OTHER INFORMATION: /label= pGlu
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"

FEATURE:

NAME/KEY: Region
LOCATION: 1..10

OTHER INFORMATION: /note= "immunomimic"
FEATURE:

NAME/KEY: Region
LOCATION: 1..17

OTHER INFORMATION: /note= "spacer"
US-08-968-466-6

Query Match 38.8%; Score 71; DB 4; Length 17;
Best Local Similarity 92.3%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSGP 14

Db 2 HWSYGLRPGSSPP 14

RESULT 15

US-08-478-546B-6
Sequence 6, Application US/08478546B

Patent No. 6303123

GENERAL INFORMATION:

APPLICANT: Grimes, Stephen

APPLICANT: Scibienski, Robert

TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent

TITLES OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormone

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dimitrios T. Drivas, Esq.

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,546B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/188,223

FILING DATE: 27-JAN-1994

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified site
LOCATION: 1
OTHER INFORMATION: /note= Xaa
OTHER INFORMATION: /note= "pyroglutamic acid"
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..17
OTHER INFORMATION: /note= "spacer"
US-08-478-546B-6

Query Match 38.8%; Score 71; DB 4; Length 17;
Best Local Similarity 92.3%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSGP 14

Db 2 HWSYGLRPGSSPP 14

Search completed: October 10, 2002, 22:48:33
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:48:39 ; Search time 16 Seconds
(without alignments)
282.262 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSGPSLKLLS.....HRLEGVPSLHWSYGLRPX 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 10944

Minimum DB seq length: 0

Maximum DB seq length: 47

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	23.4	10	1 RHPGG	gonadoliberin - pi
2	58	23.4	10	1 RSHG	gonadoliberin - Sh
3	54	21.8	10	1 RHAQ1	gonadoliberin I -
4	50	20.2	44	2 S28770	hypothetical prote
5	45	18.1	10	2 A21114	gonadoliberin - ch
6	40	16.1	10	1 RHAQ2	gonadoliberin II -
7	40	16.1	10	1 A61126	gonadoliberin - sp
8	40	16.1	10	2 A46030	gonadoliberin I -
9	40	16.1	10	2 B46030	gonadoliberin II -
10	38	15.3	21	2 A60225	pyruvate dehydroge
11	37	14.9	10	2 A49187	gonadotropin-relea
12	37	14.9	36	2 S42591	dnaA protein - Str
13	36.5	14.7	45	2 T52274	R2R3-MVB transcrip
14	36	14.5	32	2 S30756	genome polyprotein
15	36	14.5	47	2 S38586	NADH dehydrogenase
16	36	14.5	47	2 G82712	hypothetical prote
17	35	14.1	42	2 B39880	streptomycin/spect
18	34.5	13.9	35	2 C82330	hypothetical prote
19	34	13.7	29	2 S10061	Ig heavy chain (cl
20	34	13.7	39	2 H86068	partial repeat of
21	33	13.3	20	2 P50188	superoxide dismuta
22	33	13.3	36	2 D87544	hypothetical prote
23	32	12.9	25	2 S29284	hydrogenase (EC 1.
24	32	12.9	36	2 A38729	pyruvate decarboxy
25	32	12.9	40	2 B27740	gas-vesicle protei
26	32	12.9	42	2 E81911	hypothetical prote
27	32	12.9	46	2 I52716	gene neu protein -
28	32	12.9	46	2 S58637	hypothetical prote
29	31.5	12.7	40	2 S22471	hypothetical prote

30	31.5	12.7	46	2 PD0023	protein kinase (EC
31	31	12.5	21	2 C39543	collagen alpha 3(I
32	31	12.5	35	2 S44360	integrin beta-3 -
33	31	12.5	40	2 T07478	probable photosyst
34	31	12.5	43	2 S13581	collagen alpha 1(I
35	30.5	12.3	18	2 S29379	sorbitol dehydroge
36	30.5	12.3	33	2 PH1736	Ig heavy chain V r
37	30.5	12.3	34	2 PH1746	Ig heavy chain V r
38	30.5	12.3	34	2 PH1747	Ig heavy chain V r
39	30.5	12.3	34	2 PH1749	Ig heavy chain V r
40	30.5	12.3	36	2 PH1751	Ig heavy chain V r
41	30.5	12.3	36	2 PH1752	Ig heavy chain V r
42	30.5	12.3	36	2 PH1753	Ig heavy chain V r
43	30.5	12.3	39	2 G82613	hypothetical prote
44	30.5	12.3	42	2 T04383	zein protein - mai
45	30.5	12.3	46	2 A48542	CRF-related diuret

ALIGNMENTS

RESULT 1

RHPGG

gonadoliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Note: the synthetic and natural hormones have the same physicochemical and biologic

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.4%; Score 58; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.35;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10

RESULT 2

RHSHG

gonadoliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.J

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fr

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

A;Residues: 1-10 <BUR>
A;Note: the natural and synthetic hormones have the same biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormones
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.4%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 3
RHAQ1
gonadoliberin I - American alligator
N;Alternate names: gonadotropin-releasing hormone I
C;Species: Alligator mississippiensis (American alligator)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C;Accession: A60066
R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regul. Pept. 33, 105-116, 1991
A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligators
A;Reference number: A60066; MUID:91352338
A;Accession: A60066
A;Molecule type: protein
A;Residues: 1-10 <LOV>
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 21.8%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 4
S28770
hypothetical protein 44 (psb1 3' region) - barley chloroplast
C;Species: chloroplast Hordeum vulgare (barley)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C;Accession: S28770
R;Sexton, T.B.; Jones, J.T.; Mullet, J.E.
Curr. Genet. 17, 445-454, 1990
A;Title: Sequence and transcriptional analysis of the barley cDNA region upstream of psbA
A;Reference number: S28770
A;Accession: S28770
A;Molecule type: DNA
A;Residues: 1-44 <SEX>
A;Cross-references: EMBL:X52765; NID:g11601; PIDN:CAA36977.1; PID:g11607
C;Genetics:
A;Genome: chloroplast
C;Keywords: chloroplast

Query Match 20.2%; Score 50; DB 2; Length 44;
Best Local Similarity 43.5%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 3 WSYGLRPGSGPSLKLSEIKGV 25
| | | | | : | | | :
Db 19 WRPGRFPGSLDKNPRIKSEIKNI 41

RESULT 5

A21114
gonadoliberin - chum salmon
C;Species: Oncorhynchus keta (chum salmon)
C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C;Accession: A21114
R;Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A;Title: Characterization of a teleost gonadotropin-releasing hormone.
A;Reference number: A21114; MUID:83195140
A;Accession: A21114
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <SHE>

Query Match 18.1%; Score 45; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 6
RHAQ2
gonadoliberin II - American alligator
N;Alternate names: gonadotropin-releasing hormone II
C;Species: Alligator mississippiensis (American alligator)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C;Accession: B60066
R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regul. Pept. 33, 105-116, 1991
A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligators
A;Reference number: A60066; MUID:91352338
A;Accession: B60066
A;Molecule type: protein
A;Residues: 1-10 <LOV>
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 16.1%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 7
A61126
gonadoliberin - spotted ratfish
N;Alternate names: gonadotropin-releasing hormone
C;Species: Hydrophagus colliet (spotted ratfish)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C;Accession: A61126
R;Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.
Gen. Comp. Endocrinol. 82, 152-161, 1991
A;Title: Primary structure of gonadotropin-releasing hormone from the brain of a holocephalus
A;Reference number: A61126; MUID:91340067
A;Accession: A61126
A;Molecule type: protein
A;Residues: 1-10 <LOV>
A;Experimental source: brain
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 16.1%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 59;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||

Db 2 HWSHGWPY 10

RESULT 8

A46030

gonadoliberin I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 16.1%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||

Db 2 HWSHGWPY 10

RESULT 9

B46030

gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 16.1%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||

Db 2 HWSHGWPY 10

RESULT 10

A60225

pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) alpha chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 15-Oct-1999
C:Accession: A60225
R:Lawson, R.; Aitken, A.; Yeaman, S. J.
Biochem. Soc. Trans. 11, 298-299, 1983
A:Title: Primary sequence of the N-terminal region of the alpha-subunit of pyruvate dehy
A:Reference number: A60225
A:Accession: A60225
A>Status: preliminary

A:Molecule type: protein
A:Residues: 1-21 <LAW>
C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate
C:Keywords: oxidoreductase

Query Match 15.3%; Score 38; DB 2; Length 21;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 21 EIKGVIVHRLEGVEGP 36
|||||

Db 8 EIKKCDLHRL--EGP 21

RESULT 11

A49187

gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A:Title: Primary structure and biological activity of a third gonadotropin-releasing
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A>Note: sequence extracted from NCBI backbone (NCBIF:126381)

Query Match 14.9%; Score 37; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||

Db 2 HWSHDWKP 10

RESULT 12

S42591

dnaA protein - Streptomyces lividans (fragment)
C:Species: Streptomyces lividans
C:Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 28-May-1999
C:Accession: S42591
R:Zakrzewska-Czerwinska, J.; Nardmann, J.; Schrepf, H.
Mol. Gen. Genet. 242, 440-447, 1994
A:Title: Inducible transcription of the dnaA gene from Streptomyces lividans 66.
A:Reference number: S42591; MUID:94166753
A:Accession: S42591
A:Molecule type: DNA
A:Residues: 1-36 <ZAK>
A:Cross-references: GB:S69180; NID:g545764; PIDN:AAB30114.1; PID:g545765
A:Experimental source: strain 66
C:Keywords: DNA binding

Query Match 14.9%; Score 37; DB 2; Length 36;
Best Local Similarity 37.5%; Pred. No. 5.6e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 17 KLLSEIKGVIVHRLEGVEGSLHW 40
|||||

Db 18 QLLGEGRG-----QGVESKDEHW 35

RESULT 13

T52274

R2R3-MYB transcription factor [imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T52274
R:Romero, I.; Fuentes, A.; Benito, M.J.; Malpica, J.; Leyva, A.; Paz-Ares, J.

submitted to the EMBL Data Library, May 1997

A:Description: One hundred R2R3-MYB genes in the genome of Arabidopsis thaliana.

A:Reference number: Z25968

A:Accession: T52274

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-45 <ROM>

A:Cross-references: EMBL:Z95738; PIDN:CAB09170.1

A:Experimental source: ecotype Landsberg erecta; whole plants

C:Genetics:

A:Note: AUMYB10

Query Match 14.7%; Score 36.5; DB 2; Length 45;

Best Local Similarity 31.8%; Pred. No. 8.2e+02;

Matches 14; Conservative 3; Mismatches 10; Indels 17; Gaps 3;

QY 3 WSYGLRPGSGPSLK---LLSEIKGVIVHRLEGVPSLHWSYG 43

Db 3 WYNYLRPG-----LKRGNFTKEEDTIIH-----LHQAYG 32

RESULT 14

S30756

genome polyprotein - foot-and-mouth disease virus Asia (strain Asia 1) (fragment)

C:Species: Aphthovirus Asia (foot-and-mouth disease virus Asia)

C:Date: 03-May-1994 #sequence_revision 14-Sep-1994 #text_change 26-Aug-1999

C:Accession: S30756

R:Sangar, D.V.; Newton, S.E.; Rowlands, D.J.; Clarke, B.E.

Nucleic Acids Res. 15, 3305-3315, 1987

A:Title: All foot and mouth disease virus serotypes initiate protein synthesis at two se

A:Reference number: S30753; MUID:87203363

A:Accession: S30756

A:Molecule type: genomic RNA

A:Residues: 1-32 <GAN>

A:Cross-references: EMBL:M31578; NID:g210495; PIDN:AAA42658.1; PID:g210496

A:Note: the authors translated the codon CAR for residue 5 as Asp

C:Superfamily: foot-and-mouth disease virus genome polyprotein

C:Keywords: alternative initiators; polyprotein

F:1-32/Product: protein Lab (fragment) #status predicted <LAB>

F:29-32/Product: protein Lb (fragment) #status predicted <Lb1>

Query Match 14.5%; Score 36; DB 2; Length 32;

Best Local Similarity 42.1%; Pred. No. 6.5e+02;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 14 PSLKLLSEIKGVIVHRLEG 32

Db 9 PLLYALREIKALFLSRTQG 27

RESULT 15

S38586

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) 42.5K chain - potato mitochondrion (fragmen

N:Alternate names: complex 1 dehydrogenase 42.5K chain; NADH-ubiquinone oxidoreductase 4

C:Species: mitochondrion Solanum tuberosum (potato)

C:Date: 03-Feb-1994 #sequence_revision 03-Dec-1999 #text_change 10-Dec-1999

C:Accession: S46439; A49732; S38586

R:Gaebler, L.; Herz, U.; Liddell, A.; Leaver, C.J.; Schroeder, W.; Brennicke, A.; Grohma

Mol. Gen. Genet. 244, 33-40, 1994

A:Title: The 42.5 kDa subunit of the NADH: ubiquinone oxidoreductase (complex I) is in p

A:Reference number: S46439; MUID:94316190

A:Accession: S46439

A:Molecule type: DNA

A:Residues: 1-12,'S',14-25,'S',27,'S',29-47 <GA2>

A:Cross-references: EMBL:X75579; NID:g414717; PIDN:CAA53256.1; PID:g414718

A:Note: 13-Leu, 26-Leu, and 28-Leu are due to RNA editing

R:Herz, U.; Schroeder, W.; Liddell, A.; Leaver, C.J.; Brennicke, A.; Grohmann, L.

J. Biol. Chem. 269, 2263-2269, 1994

A:Title: Purification of the NADH:ubiquinone oxidoreductase (complex I) of the respirato

A:Reference number: A49732; MUID:94124587

A:Accession: A49732

A:Molecule type: protein

A:Residues: 2-34,'X',36-37 <HER>

C:Genetics:

A:Gene: nad7

A:Genome: mitochondrion

C:Superfamily: NADH dehydrogenase (ubiquinone) 49K protein

C:Keywords: electron transfer; hydrogen ion transport; mitochondrial inner membrane;
F:2-47/Product: NADH dehydrogenase (ubiquinone) 42.5K chain (fragment) #status exper-

Query Match 14.5%; Score 36; DB 2; Length 47;

Best Local Similarity 31.2%; Pred. No. 1e+03;

Matches 10; Conservative 7; Mismatches 11; Indels 4; Gaps 1;

QY 4 SYGLRPGSGGPS----LKLLSEIKGVIVHRLE 31

Db 10 NFTLNFGPQHPAAHGVLRLVLEMGGEVERAE 41

Search completed: October 10, 2002, 22:53:18

Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:48:04 ; Search time 13 Seconds
(without alignments)
139.986 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSSGSLKLLS.....HRLEGVEGPSLHWSYGLRPX 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 3384

Minimum DB seq length: 0

Maximum DB seq length: 47

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	54	21.8	10	1 GONL_ALLMI	P37041 alligator m
2	47	19.0	10	1 GONL_CLUPA	P81749 clupea pall
3	45	18.1	10	1 GON3_ONCKE	P20367 oncorhynch
4	40	16.1	10	1 GON2_CHICK	P37043 gallus gall
5	40	16.1	10	1 GONL_SQUAC	P27429 squalus aca
6	37	14.9	10	1 GON3_PETWA	P30948 petromyzon
7	35	14.1	39	1 ABRE_BOMPA	P81463 bombus pasc
8	33	13.3	38	1 MUTS_YEREN	P47763 yersinia en
9	33	13.3	43	1 BAGE_HUMAN	Q13072 homo sapien
10	33	13.3	43	1 PIV6_ADEB2	O95627 bovine aden
11	32	12.9	10	1 GONL_CHEPR	P80677 chelyosoma
12	32	12.9	26	1 PCW4_PACGO	P82426 pachycondyl
13	31	12.5	24	1 CR31_LITCI	P81851 litorea cit
14	31	12.5	40	1 PSBJ_PINTH	P41618 pinus thunb
15	31	12.5	43	1 PSBN_MESVI	Q9muu5 mesostigma
16	30.5	12.3	46	1 DIUH_PERAM	P41538 periplaneta
17	30.5	12.3	47	1 DIU2_TENMO	P56619 tenebrio mo
18	30	12.1	32	1 FER_FORCR	P18821 porphyridiu
19	30	12.1	33	1 FER_PORAE	P18820 porphyridiu
20	30	12.1	35	1 CECB_ANTPE	P01509 antheraea p
21	30	12.1	44	1 COAB_BPXF	P03622 bacterioph
22	29.5	11.9	24	1 COXJ_SHEEP	O9tr30 ovis aries
23	29	11.7	17	1 PC24_BRANA	P81097 brassica na
24	29	11.7	24	1 FRB4_LITIN	P82023 litorea inf
25	29	11.7	30	1 AMPT_BACST	P00728 bacillus st
26	29	11.7	30	1 HETA_RADMA	P58691 radianthus
27	29	11.7	31	1 SODC_STRHE	P81163 striga herm
28	29	11.7	34	1 Z33B_HUMAN	O06731 homo sapien
29	29	11.7	36	1 PETM_SYNY3	P74810 synechocyst
30	29	11.7	40	1 PSBJ_WARFO	P12188 marchantia
31	29	11.7	41	1 BAXC_HUMAN	Q07815 homo sapien
32	29	11.7	44	1 RLA2_RABIT	P19943 oryctolagus
33	29	11.7	44	1 Y723_TREPA	O83711 treponema p

34	29	11.7	46	1 LHBI_RHOPA	P35106 rhodopseudo
35	29	11.7	47	1 TXA2_ANESU	P01528 anemonia su
36	28.5	11.5	22	1 AOPA_MOUSE	Q64133 mus musculus
37	28.5	11.5	24	1 RS13_THETH	P80377 thermus aqu
38	28	11.3	10	1 GON2_CHEPR	P80678 chelyosoma
39	28	11.3	24	1 SODC_RANCA	P23417 rana catesb
40	28	11.3	28	1 PETL_CYAPA	P48102 cyanophora
41	28	11.3	36	1 NUCM_SOLTU	P80264 solanum tub
42	28	11.3	40	1 ALBI_TRASC	P81188 trachemys s
43	28	11.3	43	1 PSBN_CYAPA	P48108 cyanophora
44	28	11.3	43	1 PSBN_PORPU	P51324 porphyra pu
45	28	11.3	44	1 BGAL_PINPS	P81669 pinus pinas

ALIGNMENTS

RESULT 1

GONL_ALLMI
ID GONL_ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amigaion; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 21.8%; Score 54; DB 1; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.3; Mismatches 1; Indels 0; Gaps 0;

Matches 8; Conservative 1;

QY 2 HWSYGLRPG 10

Db 2 HWSYGLQPG 10

|||||

RESULT 2

GONL_CLUPA
ID GONL_CLUPA STANDARD; PRT; 10 AA.
AC P81749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH)
DE (Luliberin I).
GN GNRHI.
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;

```

OC Clupeinae; Clupea.
OX NCBI_TaxID=30724;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RL hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;

Query Match 19.0%; Score 47; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 2.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
   |||:| |
DB 2 HWSHGLSPG 10

RESULT 3
GN3-ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GnRH-III) (LH-
DE RH III) (Luliberin III).
GN GNRH3.
OS Oncorhynchus keta (Chum salmon), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN SEQUENCE.
RC SPECIES=O.keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Elden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone."
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RL hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR PIR; A21114; A21114.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

```

```

FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

Query Match 18.1%; Score 45; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
   |||:| |
DB 2 HWSYGLWPG 10

RESULT 4
GN2-CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GnRH-II)
DE (LH-RH II) (Luliberin II).
OS Gallus gallus (Chicken).
OS Alligator mississippiensis (American alligator),
OS Squalus acanthias (Spiny dogfish),
OS Hydrolagus colliei (Spotted ratfish) (Pacific ratfish), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN SEQUENCE.
RC SPECIES=Chicken; TISSUE=Hypothalamus;
RX MEDLINE=84222059; PubMed=6427779;
RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
RA Matsuo H.;
RT "Identification of the second gonadotropin-releasing hormone in
RT chicken hypothalamus: evidence that gonadotropin secretion is
RT probably controlled by two distinct gonadotropin-releasing hormones
RT in avian species."
RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=A.mississippiensis; TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
RT from brains of the American alligator (Alligator mississippiensis)."
RL Regul. Pept. 33:105-116(1991).
RN [3]
RP SEQUENCE.
RC SPECIES=S.acanthias; TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
RT dogfish brain provides insight into GnRH evolution."
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN [4]
RP SEQUENCE.
RC SPECIES=H.colliei; TISSUE=Brain;
RX MEDLINE=91340067; PubMed=1678723;
RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
RA Lee T.;
RT "Primary structure of gonadotropin-releasing hormone from the brain
RT of a holoccephalan (ratfish: Hydrolagus colliei)."
RL Gen. Comp. Endocrinol. 82:152-161(1991).
RN [5]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

```

RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: A61126; A61126.
DR PIR: B46030; B46030.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;

Query Match 16.1%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWLP 10

RESULT 5
GONL_SQUAC STANDARD; PRT; 10 AA.
ID GONL_SQUAC
AC P27429;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
(Luliberin).
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Naamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
dogfish brain provides insight into GNRH evolution."
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;

Query Match 16.1%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWLP 10

RESULT 6
GON3_PETMA STANDARD; PRT; 10 AA.
ID GON3_PETMA
AC P30948;

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
(Luliberin III).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
releasing hormone from lamprey brain."
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237A1F5A3 CRC64;

Query Match 14.9%; Score 37; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWLP 10

RESULT 7
ABAE_BOMPA STANDARD; PRT; 39 AA.
ID ABAE_BOMPA
AC P81463;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Abaecin.
OS Bombus pascuorum.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Apoidea; Apidae; Bombus.
OX NCBI_TaxID=65598;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=97362903; PubMed=9219367;
RA Rees J.A., Moniatte M., Bulet P.;
RT "Novel antibacterial peptides isolated from a European bumblebee,
Bombus pascuorum (Hymenoptera, Apoidea).";
RL Insect Biochem. Mol. Biol. 27:413-422(1997).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE AGAINST GRAM-POSITIVE AND
GRAM-NEGATIVE BACTERIA.
CC -!- INDUCTION: BY BACTERIAL INFECTION.
CC -!- SIMILARITY: PARTIAL TO APIDAECINS AND DIPTERICINS.
KW Insect immunity; Antibiotic; Hemolymph.
SQ SEQUENCE 39 AA; 4395 MW; 52E952E25D13A097 CRC64;

Query Match 14.1%; Score 35; DB 1; Length 39;
Best Local Similarity 29.7%; Pred. No. 36e+02;
Matches 11; Conservative 1; Mismatches 15; Indels 10; Gaps 2;

QY 8 RFGSGPSKLKLSKIKGVIVHRLEGVEGPSLHWSYGL 44
DB 8 RFGQSKP-----FPSFPG-----HGPFNPKIQWYPL 34

```

RESULT 8
ID MUTS_YEREN STANDARD; PRT; 38 AA.
AC P47763;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein mutS (Fragment).
GN MUTS
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W1024 / SEROTYPE O:9;
RX MEDLINE=95247270; PubMed=7729893;
RA Iriarte M., Stainier I., Cornelis G.R.;
RT "The rps gene from Yersinia enterocolitica and its influence on
RT expression of virulence factors.";
RL Infect. Immun. 63:1840-1847(1995).
CC -!- FUNCTION: This protein is involved in the repair of mismatches in
CC DNA. It is possible that it carries out the mismatch recognition
CC step. This protein has a weak ATPase activity (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U16152; AAC43392.1; -
DR InterPro; IPR000432; MUTS_C.
DR PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; PARTIAL.
KW DNA repair; ATP-binding; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 38 AA; 4328 MW; C67BEADA9CE88D2B CRC64;

Query Match 13.3%; Score 33; DB 1; Length 38;
Best Local Similarity 32.4%; Pred. No. 6.3e+02;
Matches 11; Conservative 4; Mismatches 13; Indels 6; Gaps 1;

QY 18 LLSEIKGVIVHRLGVGEGPS-----LHWSYGLR 45
DB 2 LLNEEVSFAVEALESIDPDLSLSPRALEWIYRLK 35

RESULT 9
BAGE_HUMAN
ID BAGE_HUMAN STANDARD; PRT; 43 AA.
AC Q13072;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE B melanoma antigen (Antigen M22-BA).
GN BAGE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95202592; PubMed=7895173;
RA Boel P., Wildmann C., Sensi M.L., Brasseur R., Renauld J.-C.,
RA Coulie P., Boon T., van der Bruggen P.;
RT "BAGE: a new gene encoding an antigen recognized on human melanomas
RT by cytolytic T lymphocytes.";

Immunity 2:167-175(1995).
-!- FUNCTION: NOT KNOWN. ANTIGEN RECOGNIZED ON A MELANOMA BY
CC AUTOLOGOUS CYTOLYTIC T LYMPHOCYTES.
CC -!- TISSUE SPECIFICITY: NOT EXPRESSED IN NORMAL TISSUES, EXCEPT IN
CC TESTIS. EXPRESSED WITH SIGNIFICANT PROPORTION IN MELANOMAS, BUT
CC ALSO IN TUMORS OF VARIOUS HISTOLOGICAL ORIGINS, SUCH AS BLADDER
CC CARCINOMAS, HEAD AND NECK SQUAMOUS CELL CARCINOMAS, LUNG AND
CC BREAST CARCINOMAS. NOT EXPRESSED IN RENAL, COLORECTAL AND
CC PROSTATIC CARCINOMAS, LEUKEMIAS AND LYMPHOMAS. MORE FREQUENTLY
CC EXPRESSED IN METASTATIC MELANOMAS THAN IN PRIMARY MELANOMAS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19180; AAC50123.1; -
DR MIM; 605167; -
KW Antigen.
SQ SEQUENCE 43 AA; 4810 MW; 36F3BBCE4012F1BB CRC64;

Query Match 13.3%; Score 33; DB 1; Length 43;
Best Local Similarity 36.4%; Pred. No. 7.1e+02;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 17 KLLSEIKGVIVHRLGVGEGPSL 38
DB 18 RLMKEESPVSWSRLEPEDGTAL 39

RESULT 10
PIV6_ADEB2
ID PIV6_ADEB2 STANDARD; PRT; 43 AA.
AC Q96627;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Minor capsid protein VI precursor (Fragment).
GN PVI.
OS Bovine adenovirus type 2 (Mastadenovirus bos2).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=114429;
RN [1]
RP SEQUENCE FROM N.A.
RA Rusvai M., Harrach B., Banreivi A., Evans P., Benko M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MINOR CAPSID PROTEIN THAT MAY ACT AS A LINK BETWEEN THE
CC EXTERNAL CAPSID AND THE INTERNAL DNA-PROTEIN CORE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U41123; AAB16759.1; -
FT PROPEP 1 33 BY SIMILARITY.
FT CHAIN 34 >43 MINOR CAPSID PROTEIN VI.
FT NON_TER 43 43
SQ SEQUENCE 43 AA; 4584 MW; 10F78E9678070306 CRC64;

Query Match 13.3%; Score 33; DB 1; Length 43;
Best Local Similarity 41.2%; Pred. No. 7.1e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 30 LEGVEGPSLHWSYGLRP 46
DB 1 MEGINFSAAPRYGSRP 17
```

```

KW Antibiotic; Insect Immunity; Fungicide; Hemolysis.
SQ SEQUENCE 26 AA; 2853 MW; E361FBB1F5D4F70F CRC64;

Query Match 12.9%; Score 32; DB 1; Length 26;
Best Local Similarity 40.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 32 GVEGSLHWSYGLRP 46
DB 1 GIWGTALKGVKLLP 15

RESULT 13
CT31_LITCI STANDARD; PRT; 24 AA.
ID CT31_LITCI STANDARD; PRT; 24 AA.
AC P81851; P81852; P81853;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Citropin 3.1.2 [Contains: Citropin 3.1.1; Citropin 3.1.1].
OS Litoria citropa (Australian blue mountains tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hyliidae;
OC Litoria.
OX NCBI_TaxID=94770;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=99435977; PubMed=10504394;
RA Wegener K.L., Wabnitz P.A., Carver J.A., Bowie J.H., Chia B.C.S.,
RA Wallace J.C., Tyler M.J.;
RT "Host defence peptides from the skin glands of the Australian blue
RT mountains tree-frog Litoria citropa. Solution structure of the
RT antibacterial peptide citropin 1.1."
RL Eur. J. Biochem. 265:627-637(1999).
CC -!- TISSUE SPECIFICITY: DORSAL AND SUBMENTAL SKIN GLANDS.
FT PEPTIDE 1 24 CITROPIN 3.1.2.
FT PEPTIDE 1 23 CITROPIN 3.1.1.
FT PEPTIDE 1 22 CITROPIN 3.1.
SQ SEQUENCE 24 AA; 2614 MW; C9001E295BDD0E15D CRC64;

Query Match 12.5%; Score 31; DB 1; Length 24;
Best Local Similarity 33.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
DB 10 LKELTGGVIEGIGV 24

RESULT 14
PSBJ_PINTH STANDARD; PRT; 40 AA.
ID PSBJ_PINTH STANDARD; PRT; 40 AA.
AC P41618;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Photosystem II reaction center J protein.
GN PSBJ.
OS Pinus thunbergii (Green pine) (Japanese black pine).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3350;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95024047; PubMed=7937893;
RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
RA Suglura M.;
RT "Loss of all ndh genes as determined by sequencing the entire
RT chloroplast genome of the black pine Pinus thunbergii."
RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
CC -!- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE REACTION CENTER
```

```
CC      OF PHOTOSYSTEM II, ITS EXACT FUNCTION IS NOT YET KNOWN.
CC      -!- SIMILARITY: BELONGS TO THE PSBJ FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D17510; BAA04356.1; -
DR      Mendel; 10006; PINth;psbj;1.
DR      InterPro; IPR002682; Psbj.
DR      Pfam; PF01788; Psbj; 1.
KW      Photosynthesis; Reaction center; Chloroplast; Photosystem II.
SQ      SEQUENCE 40 AA; 4147 MW; 21AC20BE09B4FE93 CRC64;

Query Match      12.5%; Score 31; DB 1; Length 40;
Best Local Similarity 44.4%; Pred. NO. 1.2e+03;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      16 LKLLSEIKGVIVHLEGV 33
Db      10 LMLIGTVGIIVIGLGV 27

RESULT 15
PSBN_MESVI
ID      PSBN_MESVI      STANDARD;      PRT;      43 AA.
AC      Q9MUV5;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Photosystem II reaction center N protein.
GN      PSBN.
OS      Mesostigma viride.
OG      Chloroplast.
OC      Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC      Mesostigmatales; Mesostigmataceae; Mesostigma.
OX      NCBI_TaxID=41882;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NIES-296;
RX      MEDLINE=20150907; PubMed=10688199;
RA      Lemieux C., Otis C., Turmel M.;
RT      "Ancestral chloroplast genome in Mesostigma viride reveals an early
RL      branch of green plant evolution.";
RL      Nature 403:649-652(2000).
CC      -!- FUNCTION: NOT KNOWN.
CC      -!- SIMILARITY: BELONGS TO THE PSBN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF166114; AAF43796.1; -
DR      InterPro; IPR003398; PSBN.
DR      Pfam; PF02468; Psbn; 1.
KW      Photosystem II; Transmembrane; Chloroplast.
FT      TRANSMEM 6 28 POTENTIAL.
SQ      SEQUENCE 43 AA; 4840 MW; 24D4D270A6DBC3DE CRC64;

Query Match      12.5%; Score 31; DB 1; Length 43;
Best Local Similarity 50.0%; Pred. NO. 1.3e+03;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      30 LEGVEGPSLHWSYG 43
Db      10 LKLLSEIKGVIVHLEGV 33
```

Db 15 LIGLTGYSLYTSFG 28

Search completed: October 10, 2002, 22:52:41
Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:49:34 ; Search time 29 Seconds
(without alignments)
280.371 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSSGSLKLLS.....HRLGVEGSPSLHWSYGLRPX 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 27559

Minimum DB seq length: 0

Maximum DB seq length: 47

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	50	20.2	39	Q96CW4	Q96cw4 homo sapien
2	50	20.2	44	Q32439	Q32439 hordeum vul
3	45	18.1	33	Q9PT34	Q9pt34 oncorhynchu
4	45	18.1	33	Q9W7G0	Q9w7g0 oncorhynchu
5	41	16.5	44	Q9S0H9	Q9s0h9 borrelia bu
6	40	16.1	36	Q923L1	Q923l1 mus musculu
7	40	16.1	37	Q9N260	Q9n260 sus scrofa
8	38	15.3	20	Q9EQX8	Q9eqx8 mus musculu
9	38	15.3	27	Q9IIP8	Q9iip8 hepatitis c
10	38	15.3	32	Q9U340	Q9u340 caenorhabdi
11	37	14.9	36	Q53406	Q53406 streptomyce
12	37	14.9	37	Q51818	Q51818 lactobacilli
13	37	14.9	39	Q30828	Q30828 campylobact
14	37	14.9	47	Q41781	Q41781 zea mays (m
15	36.5	14.7	45	Q49752	Q49752 arabidopsis
16	36	14.5	32	Q66857	Q66857 foot-and-mo

17	36	14.5	43	4	Q96QEL	Q96qel homo sapien
18	36	14.5	47	16	Q9PE35	Q9pe35 xylella fas
19	35.5	14.3	37	11	Q9MT9	Q9mt9 mus musculu
20	35.5	14.3	44	8	Q9MR58	Q9mr58 beta vulgar
21	35.5	14.3	46	15	Q66719	Q66719 equine infe
22	35	14.1	34	11	Q9WU00	Q9wu00 mus musculu
23	35	14.1	39	2	Q45290	Q45290 corynebacte
24	35	14.1	47	2	Q9K517	Q9k517 mycobacteri
25	34.5	13.9	32	5	Q95NV7	Q95nv7 drosophila
26	34.5	13.9	35	16	Q9KUY3	Q9kuy3 vibrio chol
27	34.5	13.9	40	4	Q9EW55	Q9ew55 homo sapien
28	34	13.7	23	2	Q9AIL1	Q9aill1 magnetospir
29	34	13.7	33	2	P82583	P82583 streptococc
30	34	13.7	42	3	Q42780	Q42780 ustilago ho
31	34	13.7	45	12	Q64828	Q64828 human adeno
32	34	13.7	45	12	Q64838	Q64838 human adeno
33	33.5	13.5	44	11	Q9R259	Q9r259 mus musculu
34	33.5	13.5	45	8	Q9MKF9	Q9mkf9 spiranthes
35	33.5	13.5	45	8	Q9MKF8	Q9mkf8 spiranthes
36	33.5	13.5	45	8	Q9MKF7	Q9mkf7 spiranthes
37	33.5	13.5	45	8	Q9MKF5	Q9mkf5 spiranthes
38	33.5	13.5	45	10	Q9SAM6	Q9sam6 arabidopsis
39	33.5	13.5	47	10	Q43185	Q43185 solanum tub
40	33	13.3	31	5	Q9TWK5	Q9twk5 mytilus edu
41	33	13.3	36	16	Q9A5R6	Q9a5r6 caulobacter
42	33	13.3	39	6	Q95J55	Q95j55 pan troglod
43	33	13.3	45	10	Q49787	Q49787 arabidopsis
44	33	13.3	45	10	Q49804	Q49804 arabidopsis
45	33	13.3	47	6	Q29327	Q29327 sus scrofa

ALIGNMENTS

RESULT 1

Q96CW4 PRELIMINARY: PRT; 39 AA.
AC Q96CW4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 4.4 KDA PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE, AND RETINOBLASTOMA;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC013784; AAH13784.1; -;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 39 AA; 4350 MW; 7F552A68A2042B10 CRC64;

Query Match 20.2%; Score 50; DB 4; Length 39;

Best Local Similarity 31.7%; Pred. No. 24;
Matches 13; Conservative 6; Mismatches 12; Indels 10; Gaps 2;

QY 3 WSYGL--RPGSSGPSLKLSEIKGVIVHRLEGVGEGPSLHWS 41

Db 6 WGVGLASPG-----LDSOKLIILHRSEGFYPTSRWT 38

RESULT 2

Q32439 PRELIMINARY: PRT; 44 AA.

ID Q32439
AC Q32439;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE ORF44.

OS Hordeum vulgare (Barley).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;

RP [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=90291518; PubMed=1694111;
 RA Sexton T.B., Jones J.T., Mullet J.E.;
 RT "Sequence and transcriptional analysis of the barley ctDNA region
 RT upstream of psbD-psbC encoding trnK(UUU), rps16, trnQ(UUG), psbK,
 RT psbI, and trnS(GCU).";
 RL Curr. Genet. 17:445-454(1990).
 DR EMBL; X52765; CAA36977.1; -;
 KW Chloroplast.
 SQ SEQUENCE 44 AA; 5188 MW; 7A11B128EB8DAE4C CRC64;

Query Match 20.2%; Score 50; DB 8; Length 44;
 Best Local Similarity 43.5%; Pred. No. 27;
 Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 3 WSYGLRPGSGSLKLLSEIKGV 25
 | | | | | : | : | | :
 Db 19 WRPGFRGSLDNPKIKSETKNI 41

RESULT 3
 O9PT34 PRELIMINARY; PRT; 33 AA.
 AC O9PT34;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
 DE (LULIBERIN) (FRAGMENT).
 GN GNRH1.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99312119; PubMed=10385393;
 RA Von Schalburg K.R., Sherwood N.M.;
 RT "Regulation and expression of gonadotropin-releasing hormone gene
 RT differs in brain and gonads in rainbow trout.";
 RL Endocrinology 140:3012-3024(1999).
 RN [2]

SEQUENCE FROM N.A.
 RA von Schalburg K.R., Sherwood N.M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC -!- SIMILARITY). BELONGS TO THE GNRH FAMILY.
 DR EMBL; AF110533; AAD43461.1; -;
 DR InterPro; IPR002047; AKH.
 DR InterPro; IPR002012; GNRH.
 DR Pfam; PF00446; GNRH; 1.
 DR PROSITE; PS00256; AKH; UNKNOWN_1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Amidation; Hormone.
 FT NON_TER 33
 SQ SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;

Query Match 18.1%; Score 45; DB 13; Length 33;
 Best Local Similarity 77.8%; Pred. No. 85;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 | | | | |
 Db 25 HWSYGLWPG 33

RESULT 4
 O9W7G0 PRELIMINARY; PRT; 33 AA.
 AC O9W7G0;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
 DE (LULIBERIN) (FRAGMENT).
 GN GNRH2.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99312119; PubMed=10385393;
 RA Von Schalburg K.R., Sherwood N.M.;
 RT "Regulation and expression of gonadotropin-releasing hormone gene
 RT differs in brain and gonads in rainbow trout.";
 RL Endocrinology 140:3012-3024(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA von Schalburg K.R., Sherwood N.M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC -!- SIMILARITY). BELONGS TO THE GNRH FAMILY.
 DR EMBL; AF110993; AAD43463.1; -;
 DR InterPro; IPR002012; GNRH.
 DR Pfam; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Amidation; Hormone.
 FT NON_TER 33
 SQ SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;

Query Match 18.1%; Score 45; DB 13; Length 33;
 Best Local Similarity 77.8%; Pred. No. 85;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 | | | | |
 Db 25 HWSYGLWPG 33

RESULT 5
 O9S0H9 PRELIMINARY; PRT; 44 AA.
 AC O9S0H9;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 5.1 KDA PROTEIN.
 GN BBS28.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid cp32-3.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,
 RA White O., Dodson R., Hickey E.K., Gwinn M., Peterson J., van-Vugt R.,
 RA Palmer N., Haft D., Rosa P., Stevenson B.;
 RT "A bacterial genome in flux: The twelve linear and nine circular
 RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
 RT spirochete Borrelia burgdorferi.";
 RL Mol. Microbiol. 0:0-0(1999).
 DR EMBL; AE001576; AAF07471.1; -;
 KW Hypothetical protein; plasmid.
 SQ SEQUENCE 44 AA; 5132 MW; 2D78723CA85F53DA CRC64;

Query Match 16.5%; Score 41; DB 2; Length 44;
Best Local Similarity 34.5%; Pred. No. 3.8e+02;
Matches 10; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 11 SSGPSKLLSEIKGVIVHRLGVGSPSLH 39
I :
DB 9 STVKNKLISEVKNYKRVFLARNSTLH 37

RESULT 6

ID Q923L1 PRELIMINARY; PRT; 36 AA.
AC Q923L1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE LYSYL HYDROXYLASE 2 (FRAGMENT).
GN PLOD2.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=129/SVJ;
RC MEDLINE=21233587; PubMed=11334715;
RA Ruotsalainen H., Vanhatupa S., Tampio M., Sipila L., Valtavaara M.,
RA Myllyla R.;
RT "Complete genomic structure of mouse lysyl hydroxylase 2 and lysyl
RT hydroxylase 3/collagen glucosyltransferase.";
RL Matrix Biol. 20:137-146(2001).
DR EMBL: AY026758; AAK20117.1; -.
FT NON_TER 36
SQ SEQUENCE 36 AA; 3719 MW; 632C26A373081E41 CRC64;

Query Match 16.1%; Score 40; DB 11; Length 36;
Best Local Similarity 50.0%; Pred. No. 4e+02;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 GLRPGSSGPSKLLSEIKGV 25
I :
DB 5 GARGPLMPMLALLSWAAGL 24

RESULT 7

ID Q9N260 PRELIMINARY; PRT; 37 AA.
AC Q9N260;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MADH4 PROTEIN (FRAGMENT).
GN MADH4.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20118593; PubMed=10654432;
RA Kelly K.A., Larsen N.J., Marklund S., Rothschild M.F.;
RT "Mapping of two tumor suppressor genes in the pig.";
RL Anim. Biotechnol. 10:81-85(1999).
DR EMBL: AF120280; AAF70206.1; -.
DR HSSP: Q13485; 1YGS.
FT NON_TER 1
FT NON_TER 37
SQ SEQUENCE 37 AA; 4038 MW; EDFA3087D75E7B30 CRC64;

Query Match 16.1%; Score 40; DB 6; Length 37;
Best Local Similarity 42.1%; Pred. No. 4.2e+02;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 27 VHRLEGVEGPSLHWSYGLR 45
I :
DB 19 VHRTEAIERARLHICKGVQ 37

RESULT 8

Q9EQX8 PRELIMINARY; PRT; 20 AA.
ID Q9EQX8
AC Q9EQX8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE GLUTAMATE DECARBOXYLASE 65 (FRAGMENT).
GN GAD65.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Makinae K., Kobayashi T., Kobayashi H., Sakagami H.,
RA Kondo H., Tashiro F., Miyazaki J., Obata K., Tamura S., Tanigawa Y.;
RT "Structure of the Mouse Glutamate Decarboxylase 65 Gene and its
RT promoter.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032757; BAB20415.1; -.
FT NON_TER 20
SQ SEQUENCE 20 AA; 1989 MW; 28F544B0E4C3768D CRC64;

Query Match 15.3%; Score 38; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 WSYGLRPGSSGP 14
I :
DB 9 WSGSEDSGADP 20

RESULT 9

ID Q9IIP8 PRELIMINARY; PRT; 27 AA.
AC Q9IIP8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Alberto S.-F.;
RT "Influence of the dynamics of Hepatitis C virus quasiespecies in the
RT histological outcome of liver transplantation.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF221397; AAF77959.1; -.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2987 MW; 04A47C9B4D26C7C2 CRC64;

Query Match 15.3%; Score 38; DB 12; Length 27;
Best Local Similarity 39.1%; Pred. No. 5.3e+02;
Matches 9; Conservative 1; Mismatches 9; Indels 4; Gaps 2;

QY 24 GVIVHRLGVGSPSLHWSYGLRP 46
I :
DB 6 GTAHTFAGLH-PSFWG---RP 24

RESULT 10

Q9U340
ID Q9U340 PRELIMINARY; PRT; 32 AA.
AC Q9U340;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE INDUCING PEPTIDE PREPROTEIN.
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE W06G6.9 PROTEIN.
GN W06G6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z83129; CAB63325.1; -
SQ SEQUENCE 32 AA; 3775 MW; E49D031288C6806A CRC64;
Query Match 15.3%; Score 38; DB 5; Length 32;
Best Local Similarity 41.7%; Pred. No. 6.4e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 HWSYGLRPGSSG 13
Db 18 HWQWAAKPGEWG 29
RESULT 11
Q53406
ID Q53406 PRELIMINARY; PRT; 36 AA.
AC Q53406;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE DNaA (FRAGMENT).
GN DNaA.
OS Streptomyces lividans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1916;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=66;
RA MEDLINE=94166753; PubMed=8121399;
RA Zakrzewska-Czerwinska J., Nardmann J., Schrempf H.;
RT "Inducible transcription of the dnaA gene from Streptomyces lividans
66".
RL Mol. Genet. 242:440-447(1994).
DR EMBL; S69180; AAB30114.1; -
FT NON_TER 36
SQ SEQUENCE 36 AA; 3973 MW; 51FFB06541B70E2E CRC64;
Query Match 14.9%; Score 37; DB 2; Length 36;
Best Local Similarity 37.5%; Pred. No. 9.7e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 6; Gaps 1;
QY 17 KLLSEIKGVIVHRLRGVGGPSLHW 40
Db 18 QLLGGRG-----QGVESKDEHW 35
RESULT 12
O51818
ID O51818 PRELIMINARY; PRT; 37 AA.
AC O51818;

DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE INDUCING PEPTIDE PREPROTEIN.
GN SPPIP.
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1599;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LTH673;
RA MEDLINE=96218690; PubMed=8636023;
RA Eijsink V.G., Brurberg M.B., Middelhoven P.H., Nes I.F.;
RT "Induction of bacteriocin production in Lactobacillus sake by a
secreted peptide".
RL J. Bacteriol. 178:2232-2237(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=LTH673;
RA MEDLINE=98043507; PubMed=9383159;
RA Brurberg M.B., Nes I.F., Eijsink V.G.;
RT "Pheromone-induced production of antimicrobial peptides in
Lactobacillus".
RL Mol. Microbiol. 26:347-360(1997).
DR EMBL; AF002276; AAB93967.1; -
SQ SEQUENCE 37 AA; 4306 MW; 324A7201075F2EAD CRC64;
Query Match 14.9%; Score 37; DB 2; Length 37;
Best Local Similarity 35.0%; Pred. No. 1e+03;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 10 GSSGPSLKLLSEIKGVIVHR 29
Db 18 GMAGSSNFHKKIKOIFTHR 37
RESULT 13
O30828
ID O30828 PRELIMINARY; PRT; 39 AA.
AC O30828;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ASPARTATE KINASE (EC 2.7.2.4) (FRAGMENT).
GN ASK.
OS Campylobacter coli.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=195;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCTC 11366;
RA Linton D., Lawson A.J., Owen R.J., Stanley J.;
RT "PCR detection, identification to species level, and fingerprinting of
Campylobacter jejuni and Campylobacter coli direct from diarrheal
samples".
RL J. Clin. Microbiol. 35:0-0(1997).
DR EMBL; AF017758; AAB70301.1; -
KW Kinase; transferase.
FT NON_TER 1
SQ SEQUENCE 39 AA; 4349 MW; 6AAAE792C770D90B CRC64;
Query Match 14.9%; Score 37; DB 2; Length 39;
Best Local Similarity 48.18; Pred. No. 1.1e+03;
Matches 13; Conservative 3; Mismatches 9; Indels 2; Gaps 2;
QY 20 SEIK-GVIVHRLRG-VEGSPSLHWSYGL 44
Db 11 SEIKISMIVHEKYGLAVRALHCEYGL 37
RESULT 14

```

Matches 14:  Conservative  3:  Mismatches 10:  Indels 17:  Gaps 3:
Qy  3  WSYGLRCGSGPSLK-----LSEIKGVIVHRLGVGEGPSLHWSYG 43
      |  |||||  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   3  WYNYLRPG-----LKRGNFTKEEDTIIH-----LHOAYG 32

```

Search completed: October 10, 2002, 22:54:09
Job time : 32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:47:29 ; Search time 28 Seconds
(without alignments)
186.445 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSSGSLKLLS.....HRLGVEGPELHWSYGLRPX 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 342942

Minimum DB seq length: 0

Maximum DB seq length: 47

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	47.2	25	15 AAR62705	LHRH-containing im
2	117	47.2	42	15 AAR62708	LHRH-containing im
3	116	46.8	27	15 AAR62707	LHRH-containing im
4	116	46.8	27	21 AAY91156	MVF Th epitope/LHR
5	116	46.8	27	21 AAY68567	Peptide immunogen
6	116	46.8	45	15 AAR62721	LHRH-containing im
7	113	45.6	27	21 AAY91163	Modified MVF Th ep
8	110	44.4	31	21 AAY91175	Modified MVF Th ep
9	108	43.5	27	21 AAY91161	Modified MVF Th ep
10	108	43.5	27	21 AAY91167	Modified MVF Th ep
11	108	43.5	45	21 AAY91165	Modified MVF Th ep

12	108	43.5	45	21 AAY68573	Peptide immunogen
13	107	43.1	31	21 AAY91179	Modified MVF Th ep
14	106	42.7	28	21 AAY91158	Modified MVF Th ep
15	106	42.7	40	20 AAW67581	Synthetic chimera
16	106	42.7	40	21 AAY79986	Measles virus fusi
17	105	42.3	31	21 AAY91173	Modified MVF Th ep
18	105	42.3	31	21 AAY68582	Peptide immunogen
19	105	42.3	47	21 AAY91180	Inv epitope/modifi
20	105	42.3	47	21 AAY68583	Peptide immunogen
21	103	41.5	27	21 AAY91170	Modified MVF Th ep
22	103	41.5	27	21 AAY68575	Peptide immunogen
23	103	41.5	35	21 AAY91242	Modified MVF Th ep
24	101	40.7	40	20 AAW67582	Synthetic chimera
25	100	40.3	28	15 AAR62726	LHRH-containing im
26	100	40.3	46	15 AAR62728	LHRH-containing im
27	100	40.3	47	21 AAY91183	Inv epitope/modifi
28	100	40.3	47	21 AAY68586	Peptide immunogen
29	99	39.9	28	21 AAY91159	Modified MVF Th ep
30	96.5	38.9	42	21 AAB20865	GnRH tandem dimer
31	94.5	38.1	30	11 AAR07323	Luteinising hormon
32	94.5	38.1	40	20 AAY31183	Ubiquitin fusion p
33	94.5	38.1	40	22 AAB71954	GnRH sequence #2
34	94.5	38.1	41	20 AAY31182	Ubiquitin fusion p
35	94.5	38.1	41	22 AAB71953	GnRH sequence #1
36	94	37.9	28	21 AAY91157	Modified MVF Th ep
37	93.5	37.7	40	21 AAY68085	Cattle gonadotropi
38	92	37.1	23	21 AAB20864	GnRH tandem repeat
39	89	35.9	27	16 AAR74261	SSAL1 TH1LHRH Sy
40	87	35.1	34	17 AAW05619	MVFLTh-GG-migE2S
41	86.5	34.9	20	19 AAW47438	Antigenic peptide
42	86.5	34.9	20	20 AAY31174	Ubiquitin fusion p
43	86.5	34.9	20	20 AAY31177	Ubiquitin fusion p
44	86.5	34.9	20	20 AAY31178	Ubiquitin fusion p
45	86.5	34.9	20	20 AAY31179	Ubiquitin fusion p

ALIGNMENTS

RESULT 1

AAR62705

ID AAR62705 standard; peptide; 25 AA.

XX

AC AAR62705;

XX

DT 10-SEP-1995 (first entry)

XX

DE LHRH-containing immunogenic peptide.

XX

KW Helper T cell epitope; universal immune stimulator; invasiv; hapten;

KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;

KW androgen-dependent carcinoma; antitumour; infertility;

KW measles virus F protein.

XX

OS Synthetic.

XX

FH Key

FT Domain

FT Domain

FT Domain

XX

PN WO9425060-A.

XX

PD 10-NOV-1994.

XX

PF 28-APR-1994; 94WO-US04832.

XX

PR 27-APR-1993; 93US-0057166.

XX

PR 14-APR-1994; 94US-0229275.

PA (LADD/) LADD A E.

PA (WANG/) WANG C Y.

Location/Qualifiers

1..15

/note= "measles virus F protein helper T cell epitope"

16..25

/note= "LHRH hapten"

```

PA (ZAMB/) ZAMB T.
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
XX that suppress LHRH activity in males and females
XX Claim 8; Page 84; 213pp; English.
XX Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein haptent containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The
XX stimulator consists of (A) a promiscuous helper T cell epitope (Th)
XX which elicits an immune response to the coupled peptide in members of
XX a heterogeneous population expressing diverse HLA phenotypes, and (B)
XX an adjuvant peptide sequence from the invasin protein of Yersinia.
XX Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
XX invasin and Th domains and between the immune stimulator and haptent
XX components. When the haptent is LHRH, then optionally the invasin domain
XX can be omitted from the immune stimulator component.
XX The present sequence represents an LHRH-containing, invasin-free
XX immunogenic peptide as above which can be used as a potent vaccine for
XX treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
XX prostatic carcinoma, testicular carcinoma, endometriosis, benign
XX uterine tumours, recurrent functional ovarian cysts, (severe)
XX premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX induction of infertility.
XX Sequence 25 AA;
XX Query Match 47.2%; Score 117; DB 15; Length 25;
XX Best Local Similarity 85.7%; Pred. No. 1.3e-08;
XX Matches 24; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 19 LSEIKGVIVHRLEGVPSLHWSYGLRP 46
DB 1 LSEIKGVIVHRLEGVPSLHWSYGLRP 24

RESULT 2
AAR62708
ID AAR62708 standard; peptide; 42 AA.
XX AC AAR62708;
XX DT 10-SEP-1995 (first entry)
XX DE LHRH-containing immunogenic peptide.
XX Helper T cell epitope; universal immune stimulator; invasin; haptent;
XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
XX androgen-dependent carcinoma; antitumour; infertility;
XX measles virus F protein.
XX Synthetic.
XX Key Location/Qualifiers
XX Domain 1..15
XX Domain 16..30
XX Domain 31..42
XX Domain /note= "LHRH haptent"
XX WO9425060-A.
XX 10-NOV-1994.
XX 28-APR-1994; 94WO-US04832.
XX

```

```

PR 27-APR-1993; 93US-0057166.
XX 14-APR-1994; 94US-0229275.
XX (LADD/) LADD A E.
XX (WANG/) WANG C Y.
XX (ZAMB/) ZAMB T.
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
XX that suppress LHRH activity in males and females
XX Claim 8; Page 86; 213pp; English.
XX Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein haptent containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The
XX stimulator consists of (A) a promiscuous helper T cell epitope (Th)
XX which elicits an immune response to the coupled peptide in members of
XX a heterogeneous population expressing diverse HLA phenotypes, and (B)
XX an adjuvant peptide sequence from the invasin protein of Yersinia.
XX Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
XX invasin and Th domains and between the immune stimulator and haptent
XX components. When the haptent is LHRH, then optionally the invasin domain
XX can be omitted from the immune stimulator component.
XX The present sequence represents an LHRH-containing, invasin-free
XX immunogenic peptide as above which can be used as a potent vaccine for
XX treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
XX prostatic carcinoma, testicular carcinoma, endometriosis, benign
XX uterine tumours, recurrent functional ovarian cysts, (severe)
XX premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX induction of infertility.
XX Sequence 42 AA;
XX Query Match 47.2%; Score 117; DB 15; Length 42;
XX Best Local Similarity 82.8%; Pred. No. 2.4e-08;
XX Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 18 LLEIKGVIVHRLEGVPSLHWSYGLRP 46
DB 15 VLSEIKGVIVHRLEGVPSLHWSYGLRP 41

RESULT 3
AAR62707
ID AAR62707 standard; peptide; 27 AA.
XX AC AAR62707;
XX DT 10-SEP-1995 (first entry)
XX DE LHRH-containing immunogenic peptide.
XX Helper T cell epitope; universal immune stimulator; invasin; haptent;
XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
XX androgen-dependent carcinoma; antitumour; infertility;
XX measles virus F protein.
XX Synthetic.
XX Key Location/Qualifiers
XX Domain 1..15
XX Domain 16..27
XX Domain /note= "measles virus F protein helper T cell epitope"
XX WO9425060-A.
XX 10-NOV-1994.
XX

```



```

XX 28-APR-1994; 94WO-US04832.
XX
XX 27-APR-1993; 93US-0057166.
XX
XX 14-APR-1994; 94US-0229275.
XX
XX (LADD//) LADD A E.
XX
XX (WANG//) WANG C Y.
XX
XX (ZAMB//) ZAMB T.
XX
XX Ladd AE, Wang CY, Zamb T;
XX
XX WPI; 1994-357910/44.
XX
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
XX that suppress LHRH activity in males and females
XX
XX Claims 8, 12; Page 86; 213pp; English.
XX
XX Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein haptan containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The
XX stimulator consists of (A) a promiscuous helper T cell epitope (Th)
XX which elicits an immune response to the coupled peptide in members of
XX a heterogeneous population expressing diverse HLA phenotypes, and (B)
XX an adjuvant peptide sequence from the invasive protein of Yersinia.
XX Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
XX invasive and Th domains and between the immune stimulator and haptan
XX components. When the haptan is LHRH, then optionally the invasive domain
XX can be omitted from the immune stimulator component.
XX The present sequence represents an LHRH-containing, invasive-free
XX immunogenic peptide as above which can be used as a potent vaccine for
XX treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
XX prostatic carcinoma, testicular carcinoma, endometriosis, benign
XX uterine tumours, recurrent functional ovarian cysts, (severe)
XX premenstrual syndrome or oestrogen-dependent breast cancer, or for
XX induction of infertility.
XX This sequence is particularly preferred.
XX
XX Sequence 27 AA;
XX
Query Match 46.8%; Score 116; DB 15; Length 27;
Best Local Similarity 85.7%; Pred. No. 1.9e-08;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 19 LSEIKGVIVHRLEGVGSLSHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | |
Db 1 LSEIKGVIVHRLEGVGSLSHWSYGLRP 26
| | | | | | | | | | | | | | | | | | | | |
RESULT 4
AA911156
ID AA911156 standard; peptide; 27 AA.
XX
XX AC AA911156;
XX
XX 22-MAY-2000 (first entry)
XX
XX MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:36.
XX
XX Hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
XX luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
XX somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
XX foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;
XX Plasmodium falciparum; circumsporozoite; antimalarial; CEST;
XX cholesterol ester transport protein; anti-arteriosclerotic.
XX
XX Chimeric - Measles virus.
OS Chimeric - Rattus sp.
XX
XX W09966957-A2.

```

Query Match 46.8%; Score 116; DB 21; Length 27;

Best Local Similarity 85.7%; Pred. No. 1.9e-08;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | | |
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 5

AAY68567
ID AAY68567 standard; peptide; 27 AA.

XX AC AAY68567;

XX DT 05-MAY-2000 (first entry)

XX DE Peptide immunogen comprising a Th epitope and LHRH target antigen.

XX KW Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; chimera;
KW immunocastration.

XX OS Chimeric - Measles virus.

XX OS Chimeric - Unidentified.

XX FH Key Location/Qualifiers

FT Peptide 1..15 /note= "helper Th epitope AAY68540"

FT Peptide 16..17 /note= "spacer"

FT Peptide 18..27 /note= "LHRH antigenic epitope AAY68566"

XX PN WO966952-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13960.

XX PR 20-JUN-1998; 98US-0100414.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX PS WPI; 2000-160562/14.

XX PT New peptide immunogen containing luteinising hormone-releasing hormone
PT antigen site and helper T cell epitope, for e.g. contraception and
PT treatment of cancer

XX PS Example 1; Page 63; 102pp; English.

XX CC The present sequence represents a peptide immunogen comprising a
CC helper T cell (Th) epitope of the F protein of the Measles virus and
CC a target antigen, luteinising hormone-releasing hormone (LHRH).
CC The peptide immunogens cause induction of a specific immune response
CC to LHRH which is involved in regulation of spermatogenesis, ovulation,
CC oestrus, sexual development and secretion of sex hormones. Provision of
CC a promiscuous T helper epitope (which is functional in genetically
CC diverse subjects) provides optimum immunogenicity to the B cell
CC epitopes of the target antigen and thus high antibody titres against
CC the target antigen. The peptide immunogens of the invention are used
CC to vaccinate against mammalian LHRH, for use as (reversible)
CC contraceptive; control of hormone-dependent tumours (cancer of prostate
CC or breast, also endometriosis); to prevent boar taint (and improve meat
CC quality) and for immunocastration.

XX SQ Sequence 27 AA;

Query Match 46.8%; Score 116; DB 21; Length 27;
Best Local Similarity 85.7%; Pred. No. 1.9e-08;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 19 LSEIKGVIVHRLEGVEGPSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | | |
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 6

AAR62721
ID AAR62721 standard; peptide; 45 AA.

XX AC AAR62721;

XX DT 10-SEP-1995 (first entry)

XX DE LHRH-containing immunogenic peptide.

XX KW Helper T cell epitope; universal immune stimulator; invasive; haptens;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility;
KW measles virus F protein.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Domain 1..16 /note= "invasin domain"

FT Domain 19..33 /note= "measles virus F protein helper T cell
epitope"

FT Domain 36..45 /note= "LHRH haptens"

XX PN WO9425060-A.

XX PD 10-NOV-1994.

XX PF 28-APR-1994; 94WO-US04832.

XX PR 27-APR-1993; 93US-0057166.

XX PR 14-APR-1994; 94US-0229275.

XX PA (LADD/) LADD A E.

XX PA (WANG/) WANG C Y.

XX PA (ZAMB/) ZAMB T.

XX PI Ladd AE, Wang CY, Zamb T;

XX PS WPI; 1994-357910/44.

XX PT Immunogenic luteinising hormone releasing hormone peptide(s) -
PT that suppress LHRH activity in males and females

XX PS Claim 8; Page 88; 213pp; English.

XX CC Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein haptens containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The
CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
CC which elicits an immune response to the coupled peptide in members of
CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
CC an adjuvant peptide sequence from the invasive protein of Yersinia.
CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
CC invasive and Th domains and between the immune stimulator and haptens
CC components. When the haptens is LHRH, then optionally the invasive domain
CC can be omitted from the immune stimulator component.
CC The present sequence represents an LHRH-containing immunogenic peptide
CC as above which can be used as a potent vaccine for treating e.g.
CC prostatic hyperplasia, androgen-dependent carcinoma, prostatic
CC carcinoma, testicular carcinoma, endometriosis, benign uterine tumours,

CC recurrent functional ovarian cysts, (severe) premenstrual syndrome or
 CC oestrogen-dependent breast cancer, or for induction of infertility.

SQ Sequence 45 AA;
 Query Match 46.8%; Score 116; DB 15; Length 45;
 Best Local Similarity 85.7%; Pred. No. 3.5e-08;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 19 LSEIKGVIVHRLEGVSGPSLHWSYGLRP 46
 |||||
 Db 19 LSEIKGVIVHRLEGVSGGE--HWSYGLRP 44

RESULT 7

AAAY91163
 ID AAAY91163 standard; peptide: 27 AA.

XX
 AC AAAY91163;

XX
 DT 22-MAY-2000 (first entry)

XX
 DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:43.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW sonatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumporoite; antimalarial; CERP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.

OS Chimeric - Rattus sp.

XX WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

PA Wang CY;

XX WPI: 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus

PS Example 1; Page 80; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration; for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper

CC epitopes. Sequence AAAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAAY91122-Y91142,
 CC AAAY91226 and AAAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAAY91156-Y91196, AAAY91227 and AAAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAAY91200 is somatostatin, and AAAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAAY91223 is a Plasmodium falciparum circumporoite (CS) target
 CC antigen, and AAAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope, and may be used in a malaria vaccine. AAAY91228-Y91231 represent
 CC CERP-derived peptides and AAAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAAY91247 and AAAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAAY91248-Y91251 and
 CC AAAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAAY91198 and AAAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 27 AA;

Query Match 45.6%; Score 113; DB 21; Length 27;

Best Local Similarity 82.1%; Pred. No. 4.7e-08;

Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

OY 19 LSEIKGVIVHRLEGVSGPSLHWSYGLRP 46

Db 1 LSEIKGVIVHRLEGVSGGE--HWSYGLRP 26

RESULT 8

AAAY91175
 ID AAAY91175 standard; peptide: 31 AA.

XX
 AC AAAY91175;

XX 22-MAY-2000 (first entry)

XX Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:55.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW sonatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumporoite; antimalarial; CERP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.

OS Chimeric - Measles virus.

OS Chimeric - Rattus sp.

XX WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

PA (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY;
 XX WPI: 2000-160564/14.
 XX
 PT New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus -
 XX
 PS Example 1; Page 84; 129pp; English.
 XX
 CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes and peptide immunogens
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CEPT) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous Th epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin, and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

Sequence 31 AA;

Query Match 44.4%; Score 110; DB 21; Length 31;
 Best Local Similarity 78.6%; Pred. No. 1.4e-07;
 Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 19 _LSIEIKGVIVHRLGEGVPSLHWSYGLRP 46
 Db 3 LSEIKGVIVHRLGEGVPSLHWSYGLRP 30

RESULT 9
 AAY91161

ID AAY91161 standard; peptide; 27 AA.

AC AAY91161;

DT 22-MAY-2000 (first entry)

DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:41.

KW Promiscuous T-cell epitope; measles virus F protein; MVF;

KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;

KW cholesteryl ester transport protein; anti-arteriosclerotic.

OS Chimeric - Measles virus.

OS Chimeric - Rattus sp.

PN WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI: 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus -
 XX
 PS Example 1; Page 79; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CEPT) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th

CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 27 AA;

Query Match 43.5%; Score 108; DB 21; Length 27;

Best Local Similarity 71.4%; Pred. No. 2.2e-07;

Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVEGSLHWSYGLRP 46

:|||||:|||||:|:|:|

Db 1 LSEIKGVIVHRLEGVEGSLHWSYGLRP 26

RESULT 10

AAY91167

ID AAY91167 standard; peptide; 27 AA.

XX AC AAY91167;

XX 22-MAY-2000 (first entry)

DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:47.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;

XX hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

XX luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

XX somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

XX foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

XX Plasmodium falciparum; circumsporozoite; antimalarial; CERP;

XX cholesteryl ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.

OS Chimeric - Rattus sp.

PN WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

PA Wang CY;

PI WPI; 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus
 XX Example 1; Page 81; 129pp; English.
 XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC

CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration; for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX-- Sequence 27 AA;

Query Match 43.5%; Score 108; DB 21; Length 27;

Best Local Similarity 75.0%; Pred. No. 2.2e-07;

Matches 21; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVEGSLHWSYGLRP 46

:|||||:|||||:|:|:|

Db 1 MSEIKGVIVHRLEGVEGSLHWSYGLRP 26

RESULT 11

AAY91165

ID AAY91165 standard; peptide; 45 AA.

XX AC AAY91165;

XX 22-MAY-2000 (first entry)

DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:45.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;

KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

luteinising hormone releasing hormone; LHRH; contraceptive; anticancer; somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV; foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic; Plasmodium falciparum; circumsporozoite; antimalarial; CPTP; cholesterol ester transport protein; anti-arteriosclerotic.

Chimeric - Measles virus.
Chimeric - Rattus sp.
WO9966957-A2.
29-DEC-1999.
21-JUN-1999; 99WO-US13975.
20-JUN-1998; 98US-0100412.
(UNBI-) UNITED BIOMEDICAL INC.
Wang CY;
WPI; 2000-160564/14.
New artificial T helper cell epitope and derived immunogens with target antigenic site, for immunization against e.g. malaria, arteriosclerosis or human immune deficiency virus -
Example 1; Page 80; 129pp; English.

The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes of the invention along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response, specifically against Plasmodium falciparum, cholesterol ester transport protein (CPTP) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration); for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AAY91121 represents a promiscuous T helper epitope from the measles virus F (WVF) protein and sequences AAY91122-Y91142, from AAY91226 and AAY91243-Y91246 represent synthetic Th epitopes based on the WVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AAY91208 is a human CD4 CD22-like domain antigenic site, and AAY91209-Y90211 are MVH Th epitope/CD4 CD22 antigenic peptides which may be used to prevent HIV infection of T cells. AAY90212 is a modified version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3 antigenic peptides which may be used in the treatment of allergies. AAY91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent CPTP-derived peptides and AAY91232-Y91241 are immunogens comprising a CPTP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and

CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive protein epitope from Yersinia species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.
XX
XX
SO Sequence 45 AA;
Query Match 43.5%; Score 108; DB 21; Length 45;
Best Local Similarity 71.4%; Pred. No. 4e-07;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
QY 19 LSEIKGVIVHRLGVEGSLHWSYGLRP 46
DB 19 ISEIKGVIVHKIEGIGGE--HWSYGLRP 44
RESULT 12
AAY68573
ID AAY68573 standard; peptide; 45 AA.
XX
AC AAY68573;
XX
DT 05-MAY-2000 (first entry)
XX
DE Peptide immunogen comprising a Th epitope and LHRH target antigen.
XX
KW Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH; luteinising hormone-releasing hormone; spermatogenesis; ovulation; oestrus; sexual development; sex hormone; promiscuous T helper epitope; vaccine; contraceptive; hormone-dependent tumour; prostate cancer; breast cancer; endometriosis; boar taint; meat quality;
KW invasin domain; immunocastration.
XX
OS Synthetic.
OS Yersinia sp.
OS Measles virus.
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Peptide 1..16 /note= "invasin domain AAY68565"
FT Peptide 17..18 /note= "spacer"
FT Peptide 19..33 /note= "helper Th epitope AAY68544"
FT Peptide 34..35 /note= "spacer"
FT Peptide 36..45 /note= "LHRH antigenic epitope AAY68566"
XX WO9966952-A1.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13960.
XX
PR 20-JUN-1998; 98US-0100414.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
XX WPI; 2000-160562/14.
XX New peptide immunogen containing luteinising hormone-releasing hormone antigen site and helper T cell epitope, for e.g. contraception and treatment of cancer -
XX
PS Claim 9; Page 71; 102pp; English.
XX The present sequence represents a peptide immunogen comprising an

CC invasin domain immunostimulatory peptide of *Yersinia* sp., a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The synthetic Th epitope is derived
 CC from a structured synthetic antigen library (SSAL) designated SSAL1 Th1.
 CC SSAL Th1 is modelled after a promiscuous epitope taken from the F protein
 CC of the Measles virus. The peptide immunogens cause induction of a
 CC specific immune response to LHRH which is involved in regulation of
 CC spermatogenesis, ovulation, oestrus, sexual development and secretion
 CC of sex hormones. Provision of a promiscuous T helper epitope (which is
 CC functional in genetically diverse subjects) provides optimum
 CC immunogenicity to the B cell epitopes of the target antigen and thus
 CC high antibody titres against the target antigen. The peptide immunogens
 CC of the invention are used to vaccinate against mammalian LHRH, for use
 CC as (reversible) contraceptive; control of hormone-dependent tumours
 CC (cancer of prostate or breast, also endometriosis); to prevent boar
 CC taint (and improve meat quality) and for immunocastration.
 CC
 CC Sequence 45 AA;
 CC
 CC Query Match 43.5%; Score 108; DB 21; Length 45;
 CC Best Local Similarity 71.4%; Pred. No. 4e-07;
 CC Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
 CC
 CC QY 19 LSEIKGVIVHRLEGVEGSLHWSYGLRP 46
 CC :|||||:|||||:|:|:|
 CC Db 19 LSEIKGVIVHKIEGIGE--HWSYGLRP 44
 CC
 CC RESULT 13
 CC AAY91179
 CC ID AAY91179 standard; peptide; 31 AA.
 CC XX
 CC AC AAY91179;
 CC XX
 CC DT 22-MAY-2000 (first entry)
 CC XX
 CC DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:59.
 CC XX
 CC KW Promiscuous T-cell epitope; measles virus F protein; MVF;
 CC KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 CC KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 CC KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 CC KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 CC KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
 CC KW cholesterol ester transport protein; anti-arteriosclerotic.
 CC XX
 CC OS Chimeric - Measles virus.
 CC OS Chimeric - Rattus sp.
 CC XX
 CC PN WO9966957-A2.
 CC XX
 CC PD 29-DEC-1999.
 CC XX
 CC PF 21-JUN-1999; 98WO-US13975.
 CC XX
 CC PR 20-JUN-1998; 98US-0100412.
 CC XX
 CC PA (UNBI-) UNITED BIOMEDICAL INC.
 CC XX
 CC PI Wang CY;
 CC XX
 CC DR WPI; 2000-160564/14.
 CC XX
 CC PT New artificial T helper cell epitope and derived immunogens with target
 CC PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 CC PT or human immune deficiency virus
 CC XX
 CC PS Example 1; Page 86; 129pp; English.
 CC XX
 CC XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,

CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91126 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CETP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 CC
 CC XX Sequence 31 AA;
 CC SQ
 CC
 CC Query Match 43.1%; Score 107; DB 21; Length 31;
 CC Best Local Similarity 75.0%; Pred. No. 3.5e-07;
 CC Matches 21; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 CC
 CC QY 19 LSEIKGVIVHRLEGVEGSLHWSYGLRP 46
 CC :|||||:|||||:|:|:|
 CC Db 3 LSEIKGVIVHKLEGMFGGEHWSYGLRP 30
 CC
 CC RESULT 14
 CC AAY91158
 CC ID AAY91158 standard; peptide; 28 AA.
 CC XX
 CC AC AAY91158;
 CC XX
 CC DT 22-MAY-2000 (first entry)
 CC XX
 CC DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:38.
 CC XX
 CC KW Promiscuous T-cell epitope; measles virus F protein; MVF;
 CC KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 CC KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 CC KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 CC KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 CC KW cholesterol ester transport protein; anti-arteriosclerotic.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:51:24 ; search time 17 seconds
(without alignments)
67.530 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSGPSLKLLS.....HRLGVGEGPSLHWSYGLRPX 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 162809

Minimum DB seq length: 0

Maximum DB seq length: 47

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	47.2	42	1	US-08-446-692-20
2	117	47.2	42	2	US-08-488-351A-20
3	116	46.8	27	1	US-08-446-692-19
4	116	46.8	27	2	US-08-488-351A-19
5	116	46.8	27	3	US-09-100-414B-36
6	116	46.8	27	4	US-09-303-323-36
7	116	46.8	45	1	US-08-446-692-33
8	116	46.8	45	2	US-08-488-351A-33
9	113	45.6	27	3	US-09-100-414B-43
10	113	45.6	27	4	US-09-303-323-43
11	110	44.4	31	3	US-09-100-414B-55
12	110	44.4	31	4	US-09-303-323-55
13	108	43.5	27	3	US-09-100-414B-41
14	108	43.5	27	3	US-09-100-414B-47
15	108	43.5	27	4	US-09-303-323-41
16	108	43.5	27	4	US-09-303-323-47
17	108	43.5	45	3	US-09-100-414B-45
18	108	43.5	45	4	US-09-303-323-45
19	107	43.1	31	3	US-09-100-414B-59
20	107	43.1	31	4	US-09-303-323-59
21	106	42.7	28	3	US-09-100-414B-38
22	106	42.7	28	4	US-09-303-323-38
23	106	42.7	40	2	US-08-460-502-10
24	105	42.3	31	3	US-09-100-414B-53
25	105	42.3	31	4	US-09-303-323-53
26	105	42.3	47	3	US-09-100-414B-60
27	105	42.3	47	4	US-09-303-323-60

28	103.5	41.7	25	1	US-08-446-692-17	Sequence 17, Appl
29	103.5	41.7	25	2	US-08-488-351A-17	Sequence 17, Appl
30	103	41.5	27	3	US-09-100-414B-50	Sequence 50, Appl
31	103	41.5	27	4	US-09-303-323-50	Sequence 50, Appl
32	103	41.5	35	3	US-09-100-414B-80	Sequence 80, Appl
33	103	41.5	35	4	US-09-303-323-80	Sequence 80, Appl
34	101	40.7	40	2	US-08-460-502-11	Sequence 11, Appl
35	100	40.3	28	1	US-08-446-692-38	Sequence 38, Appl
36	100	40.3	28	2	US-08-488-351A-38	Sequence 38, Appl
37	100	40.3	46	1	US-08-446-692-40	Sequence 40, Appl
38	100	40.3	46	2	US-08-488-351A-40	Sequence 40, Appl
39	100	40.3	47	3	US-09-100-414B-63	Sequence 63, Appl
40	100	40.3	47	4	US-09-303-323-63	Sequence 63, Appl
41	99	39.9	28	3	US-09-100-414B-39	Sequence 39, Appl
42	99	39.9	28	4	US-09-303-323-39	Sequence 39, Appl
43	94.5	38.1	40	4	US-09-026-276-35	Sequence 35, Appl
44	94.5	38.1	41	4	US-09-026-276-34	Sequence 34, Appl
45	94	37.9	28	3	US-09-100-414B-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-08-446-692-20
; Sequence 20, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-20

Query Match 47.2%; Score 117; DB 1; Length 42;
Best Local Similarity 82.8%; Pred. No. 4.9e-09;
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 18 LLSEIKGVIVHRLGVGEGPSLHWSYGLRP 46

Db 15 VLSEIKGVIVHRLGVGEGPSLHWSYGLRP 41

```
RESULT 2
US-08-488-351A-20
; Sequence 20, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-20
Query Match 47.2%; Score 117; DB 2; Length 42;
Best Local Similarity 82.8%; Pred. No. 4.9e-09;
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 18 LLSKIGKGVIVHRLEGVGSLSHWSYGLRP 46
; :| | | | | | | | | | | | | | | | | | | | |
Db 15 VLSEIKGVIVHRLEGVGGE--HWSYGLRP 41
; :| | | | | | | | | | | | | | | | | | | | |

RESULT 3
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
```

```
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-19
Query Match 46.8%; Score 116; DB 1; Length 27;
Best Local Similarity 85.7%; Pred. No. 3.9e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGSLSHWSYGLRP 46
; :| | | | | | | | | | | | | | | | | | | | |
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26
; :| | | | | | | | | | | | | | | | | | | | |

RESULT 4
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/229,275
;; FILING DATE: 14-APR-1994
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/057,166
;; FILING DATE: 27-APR-1992
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maria C.H. Lin
;; REGISTRATION NUMBER: 29,323
;; REFERENCE/DOCKET NUMBER: 1151-4146 US2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)415-8745
;; TELEFAX: (516)751-6849
;; INFORMATION FOR SEQ ID NO: 19:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-488-351A-19

Query Match 46.8%; Score 116; DB 2; Length 27;
Best Local Similarity 85.7%; Pred. No. 3.9e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46
|||||
DB 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 5
US-09-100-414B-36
;; Sequence 36, Application US/09100414B
;; Patent No. 6025468
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Chang Yi
;; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
;; TITLE OF INVENTION: IMMUNOGENS
;; NUMBER OF SEQUENCES: 106
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Morgan & Finnegan, L.L.P.
;; STREET: 345 Park Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10154-0054
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC Windows
;; SOFTWARE: Word 97
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/100,414B
;; FILING DATE: 20-JUNE-1998
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maria H. Lin
;; REGISTRATION NUMBER: 29,323
;; REFERENCE/DOCKET NUMBER: 1151-4157
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-758-4800
;; TELEFAX: 212-751-6849
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-100-414B-36

Query Match 46.8%; Score 116; DB 3; Length 27;

Best Local Similarity 85.7%; Pred. No. 3.9e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46
|||||
DB 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 6
US-09-303-323-36
;; Sequence 36, Application US/09303323
;; Patent No. 6228987
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Chang Yi
;; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
;; TITLE OF INVENTION: IMMUNOGENS
;; NUMBER OF SEQUENCES: 106
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Morgan & Finnegan, L.L.P.
;; STREET: 345 Park Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10154-0054
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC Windows
;; SOFTWARE: Word 97
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/303,323
;; FILING DATE: 30-APR-1999
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/100,414
;; FILING DATE: 20-JUNE-1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maria H. Lin
;; REGISTRATION NUMBER: 29,323
;; REFERENCE/DOCKET NUMBER: 1151-4157
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-758-4800
;; TELEFAX: 212-751-6849
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-303-323-36

Query Match 46.8%; Score 116; DB 4; Length 27;
Best Local Similarity 85.7%; Pred. No. 3.9e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46
|||||
DB 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 7
US-08-446-692-33
;; Sequence 33, Application US/08446692
;; Patent No. 5759551
;; GENERAL INFORMATION:
;; APPLICANT: Ladd, Anna
;; APPLICANT: Wang, Chang Yi
;; APPLICANT: Zamb, Timothy
;; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
;; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
;; NUMBER OF SEQUENCES: 114
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Maria C.H. Lin

```
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-33

Query Match 46.8%; Score 116; DB 1; Length 45;
Best Local Similarity 85.7%; Pred. No. 7.3e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | |
Db 19 LSEIKGVIVHRLEGVGG--HWSYGLRP 44

RESULT 8
US-08-488-351A-33
; Sequence 33, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulato
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-33

Query Match 46.8%; Score 116; DB 2; Length 45;
Best Local Similarity 85.7%; Pred. No. 7.3e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | |
Db 19 LSEIKGVIVHRLEGVGG--HWSYGLRP 44

RESULT 9
US-09-100-414B-43
; Sequence 43, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-43

Query Match 45.6%; Score 113; DB 3; Length 27;
Best Local Similarity 82.1%; Pred. No. 9.8e-09;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
```

Db 1 LSEIKGVIVHLEGVGGE--HWSYGLRP 26
|||||

RESULT 10

US-09-303-323-43
; Sequence 43, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-43

Query Match 45.6%; Score 113; DB 4; Length 27;
Best Local Similarity 82.1%; Pred. No. 9.8e-09;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHLEGVGSPSLHWSYGLRP 46
|||||

Db 1 LSEIKGVIVHLEGVGGE--HWSYGLRP 26

RESULT 11

US-09-100-414B-55
; Sequence 55, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-55

Query Match 44.4%; Score 110; DB 3; Length 31;
Best Local Similarity 78.6%; Pred. No. 2.9e-08;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHLEGVGSPSLHWSYGLRP 46
|||||

Db 3 LSEIKGVIVHLEGVLFGEHWSYGLRP 30

RESULT 12

US-09-303-323-55
; Sequence 55, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

? NO
?
? GENERAL INFORMATION:
? APPLICANT: wang, Chang Yi
? TITLE OF INVENTION: NOVEL LHRH PEPTIDE
? NUMBER OF SEQUENCES: 106
? CO-INVENTOR NAMES:
? ADDRESS: Morgan & Finnegan, L.L.P.
? STREET: 345 park Avenue
? CITY: New York
?
```

4
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-41

Query Match 43.5%; Score 108; DB 4; Length 27;
Best Local Similarity 71.4%; Pred. NO. 4.6e-08;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLGVEGSPSLHWSYGLRP 46
:|||||:|||||:|||||:|||||
Db 1 ISEIKGVIVHKIEGIGGE--HWSYGLRP 26

Search completed: October 10, 2002, 22:54:44
Job time : 18 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:55:45 ; Search time 17 Seconds
(without alignments)
282.616 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHSYGLRPGSSGSPLENNF.....VSASHLEGPGLHWSYGLRPG 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 11821

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	21.2	10	1 RHPGG	gonadoliberin - pi
2	58	21.2	10	1 RHSHG	gonadoliberin - sh
3	54	19.7	10	1 RHAQ1	gonadoliberin I -
4	45	16.4	10	2 A21114	gonadoliberin - ch
5	41.5	15.1	48	2 PN0486	cell division cycl
6	40	14.6	10	1 RHAQ2	gonadoliberin II -
7	40	14.6	10	1 A61126	gonadoliberin - sp
8	40	14.6	10	2 A46030	gonadoliberin I -
9	40	14.6	10	2 B46030	gonadoliberin II -
10	40	14.6	30	2 S63531	gonadoliberin I -
11	37	13.5	10	2 A49187	gonadoliberin - sp
12	37	13.5	27	2 PN0584	gonadoliberin - sp
13	36	13.1	40	2 S12207	gonadoliberin - sp
14	35.5	13.0	36	2 PH1751	gonadoliberin - sp
15	35	12.8	27	2 PN0583	gonadoliberin - sp
16	35	12.8	32	2 F71357	gonadoliberin - sp
17	35	12.8	38	2 A82450	gonadoliberin - sp
18	35	12.8	44	2 S28770	gonadoliberin - sp
19	34	12.4	35	2 C25159	gonadoliberin - sp
20	34	12.4	39	2 H86068	gonadoliberin - sp
21	34	12.4	45	2 T32817	gonadoliberin - sp
22	34	12.4	48	2 T27285	gonadoliberin - sp
23	34	12.4	49	2 PX0079	gonadoliberin - sp
24	34	12.4	50	2 A60718	gonadoliberin - sp
25	34	12.4	50	2 C99799	gonadoliberin - sp
26	33.5	12.2	47	1 NAXACE	gonadoliberin - sp
27	33.5	12.2	47	2 D97948	gonadoliberin - sp
28	33.5	12.2	47	2 B25860	gonadoliberin - sp
29	33.5	12.2	49	1 NAXA	gonadoliberin - sp

30	33.5	12.2	50	2 A61149	gonadoliberin - pig
31	33.5	12.2	50	2 F71281	gonadoliberin - pig
32	33	12.0	27	2 PN0585	gonadoliberin - pig
33	33	12.0	33	2 F87516	gonadoliberin - pig
34	33	12.0	38	2 D84227	gonadoliberin - pig
35	33	12.0	44	2 F84645	gonadoliberin - pig
36	33	12.0	47	2 T01680	gonadoliberin - pig
37	33	12.0	48	2 A46232	gonadoliberin - pig
38	33	12.0	49	2 G44530	gonadoliberin - pig
39	33	12.0	50	2 S26941	gonadoliberin - pig
40	32.5	11.9	26	2 T11816	gonadoliberin - pig
41	32.5	11.9	37	2 T06571	gonadoliberin - pig
42	32.5	11.9	47	1 T2AZ	gonadoliberin - pig
43	32.5	11.9	49	2 S0503	gonadoliberin - pig
44	32	11.7	27	2 PN0586	gonadoliberin - pig
45	32	11.7	32	2 A56589	gonadoliberin - pig

ALIGNMENTS

RESULT 1

RHPGG

C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Note: the synthetic and natural hormones have the same physicochemical and biologic

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:10/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 21.2%; Score 58; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

DB 2 HWSYGLRPG 10

RESULT 2

RHSHG

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

endo-1,4-beta-xyla
hypothetical prote
tyrosine 3-monooxy
hypothetical prote
hypothetical prote
60S ribosomal prot
sepiasterin reduct
basic helix-loop-h
T-cell receptor al
Ig heavy chain V r
hypothetical prote
hypothetical prote
toxin II - snake-l
protein kinase Mpk
tyrosine 3-monooxy
galactose binding

A:Residues: 1-10 <BUR>
A>Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 21.2%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 3
RHAQ1
gonadoliberein I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 19.7%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.65;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 4
A21114
gonadoliberein - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W. Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match 16.4%; Score 45; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 8.9;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 5
PN0486
cell division cycle-2-related protein kinase (EC 2.7.1.1-) CRK5 - mouse (fragment)
C:Species: Mus musculus (house mouse)

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C:Accession: PN0486
R:Ershler, M.A.; Nagorskaya, T.V.; Visser, J.W.M.; Belyavsky, A.V. Gene 124, 305-306, 1993
A:Title: Novel CDC2-related protein kinases produced in murine hematopoietic stem cell lines
A:Reference number: PN0479; MUID:93185941
A:Accession: PN0486
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-48 <ERS>
A:Cross-references: EMBL:X64606
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: cell cycle control; phosphotransferase

Query Match 15.1%; Score 41.5; DB 2; Length 48;
Best Local Similarity 33.3%; Pred. No. 1.5e+02;
Matches 9; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Qy 5 YGLRPGSGPS-LFNFTVSEWLRVPK 30
:|: | | : | : | : | : | :
Db 22 FGLARAKSIPTKYTSNEVVTLMYRAPE 48

RESULT 6
RHAQ2
gonadoliberein II - American alligator
N:Alternate names: gonadotropin-releasing hormone II
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: B60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brain.
A:Reference number: A60066; MUID:91352338
A:Accession: B60066
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 14.6%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 7
A61126
gonadoliberein - spotted ratfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Hydrolyagus coliei (spotted ratfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C:Accession: A61126
R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T. Gen. Comp. Endocrinol. 82, 152-161, 1991
A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a holocephalus
A:Reference number: A61126; MUID:91340067
A:Accession: A61126
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Experimental source: brain
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 14.6%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 38;

Matches : 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:| ||

Db 2 HWSHGWPY 10

RESULT 8
A46030
gonadoliberin I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 14.6%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:| ||

Db 2 HWSHGWPY 10

RESULT 9
B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 14.6%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:| ||

Db 2 HWSHGWPY 10

RESULT 10
S63531
hypotheical protein I - Sulfobolus solfataricus (fragment)
C:Species: Sulfobolus solfataricus
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S63531
R:Jones, C.E.; Fleming, T.M.; Cowan, D.A.; Littlechild, J.A.; Piper, P.W.
Eur. J. Biochem. 233, 800-808, 1995
A:Title: The phosphoglycerate kinase and glyceraldehyde-3-phosphate dehydrogenase genes
expression in Escherichia coli.
A:Reference number: S63528; MUID:96085144
A:Accession: S63531

A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-30 <JON>
A:Cross-references: EMBL:X80178

Query Match 14.6%; Score 40; DB 2; Length 30;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 12 SGPSLFNNFTVSFWLR 27
| | | | | | | | | |

Db 2 SEPSYF---VSFWLR 13

RESULT 11
A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A:Title: Primary structure and biological activity of a third gonadotropin-releasing
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 13.5%; Score 37; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 91;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:| ||

Db 2 HWSHDWKP 10

RESULT 12
PN0584
tyrosine 3-monooxygenase (EC 1.14.16.2) type 3 and 4 - gorilla (fragment)
N:Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase
C:Species: Gorilla gorilla (gorilla)
C:Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000
C:Accession: PN0584
R:Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.
Biochem. Biophys. Res. Commun. 195, 158-165, 1993
A:Title: Increased heterogeneity of tyrosine hydroxylase in humans.
A:Reference number: PN0575; MUID:93371398
A:Accession: PN0584
A:Molecule type: genomic RNA
A:Residues: 1-27 <ICH>
A:Cross-references: GB:L14796
A:Experimental source: lymphocytes of peripheral blood
C:Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine bi
C:Superfamily: phenylalanine 4-monooxygenase
C:Keywords: bioppterin; monooxygenase; oxidoreductase

Query Match 13.5%; Score 37; DB 2; Length 27;
Best Local Similarity 36.7%; Pred. No. 2.8e+02;
Matches 11; Conservative 3; Mismatches 12; Indels 4; Gaps 1;

QY 10 GSGGPSLFNNFTVSFWLRPKVSASHLEGP 39
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 1 GAPGPSL-----TGSLLWPGTAAAPAASTPTP 26

RESULT 13
S12207
hypotheical protein (B2 element) - mouse
C:Species: Mus musculus (house mouse)

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 18-Jun-1993

R;Accession: S12207

submitted to MIPS, January 1991

A;Reference number: S12205

A;Accession: S12207

A;Molecule type: mRNA

A;Residues: 1-40 <OBE>

A;Cross-references: EMBL:X56974

Query Match 13.1%; Score 36; DB 2; Length 40;
Best Local Similarity 28.2%; Pred. No. 6e+02;
Matches 11; Conservative 2; Mismatches 4; Indels 22; Gaps 2;

QY 25 WLRVPKVSASHLEGP-----SLHWS 44

I:| | | | |

Db 4 WVRAPDCSS---EGPEFKSQPHGSGQSPVTRSDSLFWS 39

RESULT 14

PH1751

Ig heavy chain V region (clone NP-12-17) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C;Accession: PH1751

R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A;Title: Antigen-driven B cell differentiation in vivo.

A;Reference number: PH1675; MUID:93301607

A;Accession: PH1751

A;Molecule type: mRNA

A;Residues: 1-36 <MCH>

A;Experimental source: B cell

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

Query Match 13.0%; Score 35.5; DB 2; Length 36;
Best Local Similarity 27.3%; Pred. No. 6.1e+02;
Matches 12; Conservative 2; Mismatches 17; Indels 13; Gaps 2;

QY 3 WSYGLRPGS-SGPSLFNFTVSFWLRVPKVSASHLEGP SLHWSY 45

I:| | | | |

Db 5 WVKREPGTKKPSM-----QSTERYHYGSSCHGDY 36

RESULT 15

PN0583

tyrosine 3-monooxygenase (EC 1.14.16.2) type 3 and 4 - chimpanzee (fragment)

N;Alternate names: oxygen oxidoreductase; tetrahydropteridine, tyrosine hydroxylase

C;Species: Pan troglodytes (chimpanzee)

C>Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000

C;Accession: PN0583

R;Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.

Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A;Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A;Reference number: PN0575; MUID:93371398

A;Accession: PN0583

A;Molecule type: genomic RNA

A;Residues: 1-27 <ICH>

A;Cross-references: GB:L14790

A;Experimental source: lymphocytes of peripheral blood

C;Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis

C;Superfamily: phenylalanine 4-monooxygenase

C;Keywords: biopterin; monooxygenase; oxidoreductase

Query Match 12.8%; Score 35; DB 2; Length 27;
Best Local Similarity 36.7%; Pred. No. 5.1e+02;
Matches 11; Conservative 3; Mismatches 12; Indels 4; Gaps 1;

QY 10 GSSGPSLFNFTVSFWLRVPKVSASHLEGP 39

I:| | | | |

Db 1 GAPGPSL----TGSPWPGTAAAPAASTPSP 26

Search completed: October 10, 2002, 23:00:18

Job time : 18 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:54:15 ; Search time 13 Seconds
(without alignments)
148.921 Million cell updates/sec

Title: US-09-848-834a-18

Perfect score: 274

Sequence: 1 XHWSYGLRPGSSGSLNNF.....VSASHLEGPSLHWSYGLRPX 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 3667

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	54	19.7	10	1 GON1_ALIMI	P37041 alligator m
2	47	17.2	10	1 GON1_CLUPA	P81749 clupea pall
3	45	16.4	10	1 GON3_ONCKE	P20367 oncorhynch
4	40	14.6	10	1 GON2_CHICK	P37043 gallus gall
5	40	14.6	10	1 GONL_SQUAC	P27429 squalus aca
6	37	13.5	10	1 GON3_PETMA	P30948 petromyzon
7	35	12.8	32	1 Y169_TREPA	O83199 treponema p
8	33.5	12.2	39	1 ABAE_BOMPA	P81463 bombus pasc
9	33.5	12.2	47	1 TXAC_ATEL	P01532 anthopleura
10	33.5	12.2	48	1 TXA2_ANTFU	P10454 anthopleura
11	33.5	12.2	49	1 TXAA_ANTXA	P01530 anthopleura
12	33.5	12.2	50	1 Y799_TREPA	O83777 treponema p
13	32.5	11.9	47	1 TXA2_ANESU	P01528 anemonia su
14	32	11.7	10	1 GON1_CHEPR	P80677 chelyosoma
15	32	11.7	12	1 UR2_SCYCA	P35490 scyllorhinu
16	32	11.7	43	1 PIV6_ADEB2	Q96627 bovine aden
17	32	11.7	50	1 RL39_SCHPO	P05767 schizosacch
18	30.5	11.1	29	1 GALA_SHEEP	P31234 ovis aries
19	30.5	11.1	41	1 Y811_TREPA	O83788 treponema p
20	30.5	11.1	44	1 PSBN_CHLRE	O06480 chlamydomon
21	30.5	11.1	47	1 TXAL_ANTFU	P10453 anthopleura
22	30	10.9	31	1 SARL_HUMAN	O00631 homo sapien
23	30	10.9	44	1 BGAL_FINPS	P81669 pinus pinas
24	29.5	10.8	22	1 LANM_STRMU	P80666 streptococc
25	29.5	10.8	44	1 PSBN_CHLUV	P56326 chlorella v
26	29.5	10.8	44	1 RK32_LYCES	P36493 lycopersico
27	29.5	10.8	50	1 SPRT_RAT	P81728 rattus norv
28	29	10.6	20	1 LPP3_HUMAN	P56643 homo sapien
29	29	10.6	38	1 PSBL_ODOSI	P49514 odontella s
30	29	10.6	50	1 VG38_BPMD2	O64229 mycobacteri
31	28.5	10.4	31	1 CXD6_CONNI	P56710 conus nigro
32	28.5	10.4	36	1 NUCM_SOLTU	P80264 solanum tub
33	28.5	10.4	46	1 TXA5_ANESU	P01529 anemonia su

34	28	10.2	10	1 GON2_CHEPR	P80678 chelyosoma
35	28	10.2	18	1 AGL_EUPMA	P33889 euphorbia m
36	28	10.2	27	1 GRP_CHICK	P01295 gallus gall
37	28	10.2	31	1 ER29_BOVIN	P81623 bos taurus
38	28	10.2	43	1 CPC6_CANPG	P81585 cancer pagu
39	28	10.2	45	1 CYC6_PROHO	P81244 prochloroth
40	28	10.2	48	1 BA12_SCHCO	O02593 schizophyll
41	27.5	10.0	29	1 GALA_AMICA	P47214 amia calva
42	27.5	10.0	43	1 PSBN_MESVI	O9muv5 mesostigma
43	27.5	10.0	48	1 CANG_BUNCN	P82803 bunodosoma
44	27	9.9	18	1 SODM_MYCHA	P80582 mycobacteri
45	27	9.9	32	1 ER29_CHICK	P81628 gallus gall

ALIGNMENTS

RESULT 1

ID	GON1_ALIMI	STANDARD;	PRT;	10 AA.
AC	P37041; P20407;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)			
DE	(Luliberin I).			
OS	Alligator mississippiensis (American alligator).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Crocodylidae; Alligatorinae; Alligator.			
OX	NCBI_TaxID=8496;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Brain;			
RX	MEDLINE=91352338; PubMed=1882082;			
RA	Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,			
RA	Lance V., Swanson P., Rivier J.E., Sherwood N.M.;			
RT	"Primary structure of two forms of gonadotropin-releasing hormone			
RT	from brains of the American alligator (Alligator mississippiensis).";			
RL	Regul. Pept. 33:105-116(1991).			
CC	-!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
DR	PIR: A60066; RHA01.			
DR	InterPro: IPR002012; GNRH.			
DR	Pfam: PF00446; GNRH; 1.			
DR	PROSITE: PS00473; GNRH; 1.			
KW	Hormone; Amidation; Hypothalamus.			
FT	MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.			
FT	MOD_RES 10 10 AMIDATION.			
SQ	SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;			

Query Match 19.7%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.53;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
|||||:|
DB 2 HWSYGLQPG 10

RESULT 2

ID	GON1_CLUPA	STANDARD;	PRT;	10 AA.
AC	P81749;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH)			
DE	(Luliberin I).			
OS	Clupea pallasi (Pacific herring).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;			

OC Clupeinae; Clupea.
OX NCBI_TaxID=30724;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
FT MOD_RES 10 10
SQ SEQUENCE 10 AA: 1105 MW; 284B20B72871F5A3 CRC64;

Query Match 17.28; Score 47; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.8; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 1;

Qy 2 HWSYGLRPG 10
Db 2 HWSHGLSPG 10
|||:|:| |
|||:|:| |

RESULT 3
ID GON3_ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GnRH-III) (LH-
RH III) (Luliberin III).
GN GNRH3.
OS Oncorhynchus keta (Chum salmon), and
OC Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O. keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C. pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.

FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA: 1230 MW; 284B3233786B45A3 CRC64;

Query Match 16.48; Score 45; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 6.6; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2;

Qy 2 HWSYGLRPG 10
Db 2 HWSYGLWLPG 10
|||:|:| |
|||:|:| |

RESULT 4
ID GON2_CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GnRH-II)
(LH-RH II) (Luliberin II).
OS Gallus gallus (Chicken).
OS Alligator mississippiensis (American alligator),
OS Squatus acanthias (Spiny dogfish),
OS Hydrolagus colliei (Spotted ratfish) (Pacific ratfish), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=Chicken; TISSUE=Hypothalamus;
RX MEDLINE=84222059; PubMed=6427779;
RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
Matsuo H.;
RT "Identification of the second gonadotropin-releasing hormone in
chicken hypothalamus: evidence that gonadotropin secretion is
probably controlled by two distinct gonadotropin-releasing hormones
in avian species.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=A. mississippiensis; TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:103-116(1991).
RN [3]
RP SEQUENCE.
RC SPECIES=S. acanthias; TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
dogfish brain provides insight into GnRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN [4]
RP SEQUENCE.
RC SPECIES=H. colliei; TISSUE=Brain;
RX MEDLINE=91340067; PubMed=1678723;
RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
Lee T.;
RT "Primary structure of gonadotropin-releasing hormone from the brain
of a holocephalan (ratfish: Hydrolagus colliei).";
RL Gen. Comp. Endocrinol. 82:152-161(1991).
RN [5]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C. pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

```

RA Chang J.-P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RL hormones, including a novel form, from an ancient teleost, herring.";
RL Endocrinology 141:505-512(2000).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR; B60066; RHAQ2.
DR PIR; A61126; A61126.
DR PIR; B46030; B46030.
DR InterPro; IPR002012; GNRH.
DR PROSITE; PS00446; GNRH; 1.
DR Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;

Query Match 14.6%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWVPG 10

RESULT 5
GONL_SQUAC
ID GONL_SQUAC STANDARD; PRT; 10 AA.
AC P27429;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
DE (Luliberin).
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
RT dogfish brain provides insight into GNRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro; IPR002012; GNRH.
DR PROSITE; PS00446; GNRH; 1.
DR Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;

Query Match 14.6%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWVPG 10

RESULT 6
GON3_PETMA
ID GON3_PETMA STANDARD; PRT; 10 AA.
AC P30948;

```

```

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
DE (Luliberin III).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Iovas S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
RT releasing hormone from lamprey brain.";
RL Endocrinology 132:1125-1131(1993).
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro; IPR002012; GNRH.
DR PROSITE; PS00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237A1F5A3 CRC64;

Query Match 13.5%; Score 37; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 61;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWVPG 10

RESULT 7
Y169_TREPA
ID Y169_TREPA STANDARD; PRT; 32 AA.
AC O83199;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0169.
GN TP0169.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL; AE001201; AAC65161.1; -.
 DR TIGR; TP0169; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 32 AA; 3616 MW; C01A97771405398F CRC64;

Query Match 12.8%; Score 35; DB 1; Length 32;
 Best Local Similarity 36.1%; Pred. No. 3.6e+02;
 Matches 13; Conservative 2; Mismatches 15; Indels 6; Gaps 2;

Oy 3 WSYGLRPGSSGSLFNNFTVSWFLRVPKVSASHLEG 38
 ||||| : : : : : || : : || : : ||
 Db 2 WS-GLFPDQGTAFRAWVASARFV-----FHGEG 31

RESULT 8
 ID ABAE_BOMPA STANDARD; PRT; 39 AA.
 AC P81463; 1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Abaeicin.
 OS Bombus pascuorum.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 OX Apoidea; Apidae; Bombus.
 NCBI_TaxID=65598;
 [1]
 RN SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=97362903; PubMed=9219367;
 RA Rees J.A., Moniatte M., Bulet P.;
 RT "Novel antibacterial peptides isolated from a European bumblebee,
 Bombus pascuorum (Hymenoptera, Apoidea).";
 RL Insect Biochem. Mol. Biol. 27:413-422(1997).
 CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE AGAINST GRAM-POSITIVE AND
 CC GRAM-NEGATIVE BACTERIA.
 CC -1- INDUCTION: BY BACTERIAL INFECTION.
 CC -1- SIMILARITY: PARTIAL TO APIDAEICINS AND DIPTERICINS.
 KW Insect immunity; Antibiotic; Hemolymph.
 SQ SEQUENCE 39 AA; 4395 MW; 52E952E25D13A097 CRC64;

Query Match 12.2%; Score 33.5; DB 1; Length 39;
 Best Local Similarity 25.0%; Pred. No. 6.7e+02;
 Matches 10; Conservative 1; Mismatches 16; Indels 13; Gaps 1;

Oy 8 RPSGSGSLFNNFTVSWFLRVPKVSASHLEGSLHWSYGL 47
 ||||| : : : : : || : : || : : ||
 Db 8 RPSGSKP-----FPSPFGHGPFNPKIQWPYPL 34

RESULT 9
 ID TXAC_ANTEL STANDARD; PRT; 47 AA.
 AC P01532;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Anthopleurin C (Toxin AP-C).
 OS Anthopleura elegantissima (Sea anemone).
 OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Actiniaria;
 CC Nynanthaeae; Actiniidae; Anthopleura.
 NCBI_TaxID=6110;
 [1]
 RN SEQUENCE.
 RX MEDLINE=81090973; PubMed=6108877;
 RA Bennett C.D.;
 RL Unpublished results, cited by:
 RL Norton T.R.;
 RL Fed. Proc. 40:21-25(1981).
 CC -1- FUNCTION: BINDS SPECIFICALLY TO THE SODIUM CHANNEL, THEREBY
 CC DELAYING ITS INACTIVATION DURING SIGNAL TRANSDUCTION. THUS
 CC IT STRONGLY STIMULATES MAMMALIAN CARDIAC MUSCLE CONTRACTION.

CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE SEA ANEMONE SODIUM CHANNEL INHIBITORY
 CC TOXIN FAMILY.
 DR PIR: A01793; NAXACE.
 DR HSSP; P01530; IAHL.
 DR InterPro; IPR000693; Anemome_toxin.
 DR Pfam; PF00706; toxin_4; 1.
 KW Toxin; Neurotoxin; Sodium channel inhibitor.
 FT DISULFID 4 44 BY SIMILARITY.
 FT DISULFID 6 34 BY SIMILARITY.
 FT DISULFID 27 45 BY SIMILARITY.
 SQ SEQUENCE 47 AA; 4884 MW; BA0AFF6AADAID9CB CRC64;

Query Match 12.2%; Score 33.5; DB 1; Length 47;
 Best Local Similarity 38.5%; Pred. No. 8.2e+02;
 Matches 15; Conservative 4; Mismatches 11; Indels 9; Gaps 5;

Oy 11 SSGPSLFNNFTVSWFLRVPKVSASH---LEGPSLHW 43
 ||||| : : : : : || : : || : : ||
 Db 8 SDGPSVRGN-TLSGILWLAGCP--SGWHNCKAHGPTIGW 43

RESULT 10
 ID TXA2_ANTFU STANDARD; PRT; 48 AA.
 AC P10454;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Neurotoxin 2 (toxin AFII) (Toxin AFT-II).
 OS Anthopleura fuscoviridis (Sea anemone).
 CC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Actiniaria;
 CC Nynanthaeae; Actiniidae; Anthopleura.
 NCBI_TaxID=6111;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=87206977; PubMed=2883740;
 RA Sunahara S., Muramoto K., Tenma K., Kamiya H.;
 RT "Amino acid sequence of two sea anemone toxins from Anthopleura
 fuscoviridis.";
 RL Toxicon 25:211-219(1987).
 CC -1- FUNCTION: BINDS SPECIFICALLY TO THE SODIUM CHANNEL, THEREBY
 CC DELAYING ITS INACTIVATION DURING SIGNAL TRANSDUCTION. THUS
 CC IT STRONGLY STIMULATES MAMMALIAN CARDIAC MUSCLE CONTRACTION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE SEA ANEMONE SODIUM CHANNEL INHIBITORY
 CC TOXIN FAMILY.
 DR PIR; B25860; B25860.
 DR HSSP; P01530; IAHL.
 DR InterPro; IPR000693; Anemome_toxin.
 DR Pfam; PF00706; toxin_4; 1.
 KW Toxin; Neurotoxin; Sodium channel inhibitor.
 FT DISULFID 5 45 BY SIMILARITY.
 FT DISULFID 7 35 BY SIMILARITY.
 FT DISULFID 28 46 BY SIMILARITY.
 SQ SEQUENCE 48 AA; 4941 MW; 3CD4B66FADALD999 CRC64;

Query Match 12.2%; Score 33.5; DB 1; Length 48;
 Best Local Similarity 38.5%; Pred. No. 8.4e+02;
 Matches 13; Conservative 4; Mismatches 11; Indels 9; Gaps 5;

Oy 11 SSGPSLFNNFTVSWFLRVPKVSASH---LEGPSLHW 43
 ||||| : : : : : || : : || : : ||
 Db 9 SDGPSVRGN-TLSGILWLAGCP--SGWHNCKAHGPTIGW 44

RESULT 11
 ID TXAA_ANTXA STANDARD; PRT; 49 AA.
 AC P01530;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)


```

RC STRAIN=NICHOLS;
RA MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RA "Complete genome sequence of Treponema pallidum, the syphilis
RA spirochete";
RA Science 281:375-388(1998).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on way
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001250; AAC65768.1; -.
CC TIGR; TP0799; -.
DR Hypothetical protein; Complete proteome.
SK SEQUENCE 50 AA; 5796 MW; 6E788CB5CC9A818F CRC64;
KW
Query Match 12.2%; Score 33.5; DB 1; Length 50;
Best Local Similarity 29.3%; Pred. No. 8.7e+02;
Matches 12; Conservative 4; Mismatches 6; Indels 19; Gaps 2;
QY 24 FWLRV-PK-----VSASHLEGPSLHWSY 45
II II II :: I III I:
DB 5 FQQQVLPKAGRRKEHPVQYMPHKKEENATGLMNPNSLHTSH 45
RESULT 13
TXA2_ANESU STANDARD; PRT; 47 AA.
AC P01528;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DS Neurotoxin 2 (Toxin ATX-II) (Toxin II).
OS Anemonia sulcata (Snake-locks sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthae; Actiniidae; Anemonia.
OX NCBI_TaxID=6108;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=77003075; PubMed=9283;
RA Wunderer G., Fritz H., Wächter E., Machleidt W.;
RA "Amino-acid sequence of a coelenterate toxin: toxin II from Anemonia
RA sulcata.";
RT sulcata.;
RL Eur. J. Biochem. 68:193-198(1976).
RN [2]
RP DISULFIDE BONDS.
RC TISSUE=Venom;
RX MEDLINE=79046829; PubMed=30689;
RA Wunderer G.;
RT "The disulfide bridges of toxin II from Anemonia sulcata.";
RL Hoppe-Seyler's Z. Physiol. Chem. 359:1193-1201(1978).
CC -I- FUNCTION: BINDS SPECIFICALLY TO THE SODIUM CHANNEL, THEREBY
CC DELAYING ITS INACTIVATION DURING SIGNAL TRANSDUCTION, THUS
CC IT STRONGLY STIMULATES MAMMALIAN CARDIAC MUSCLE CONTRACTION.
CC -I- SUBCELLULAR LOCATION: secreted.
CC -I- SIMILARITY: BELONGS TO THE SEA ANEMONE SODIUM CHANNEL INHIBITORY
CC TOXIN FAMILY.
DR PIR: A01791; TZA2.
DR HSP; P01530; IAH1.
DR InterPro: IPR000693; Anemome_toxin.
DR Pfam: PF00706; toxin_4; 1.
DR KQ Toxin; Neurotoxin; Sodium channel inhibitor.

```

FT DISULFID 4 44
 FT DISULFID 6 34
 FT DISULFID 27 45
 FT VARIANT 2 2
 SQ SEQUENCE 47 AA; 4941 MW; BAQAF370FAID9CB CRC64;
 V -> I (IN 50% OF THE MOLECULES).

Query Match 11.9%; Score 32.5; DB 1; Length 47;
 Best Local Similarity 38.5%; Pred. No. 1.1e+03;
 Matches 15; Conservative 4; Mismatches 11; Indels 9; Gaps 5;

Qy 11 SSGPSLNNFTVS--FWLR-VPKVSASH---LEGPSLHW 43
 | ||| : | : | | | | | | | | | | | | | | | |
 Db 8 SDGPSVRGN-TLSGIWLAGCP--SGWHNCKKHGPTIGW 43

RESULT 14

ID GONI_CHEPR STANDARD; PRT; 10 AA.
 AC P80677;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I)
 DE (Luliberin I).
 OS Chelyosoma productum.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
 OC Corellidae; Chelyosoma.
 OX NCBI_TaxID=71177;
 RN [1]

RP SEQUENCE.
 RX MEDLINE=96413669; PubMed=8816823;
 RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
 RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
 RT "Two new forms of gonadotropin-releasing hormone in a protochordate
 and the evolutionary implications.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
 CC THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
 CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
 CC -1- MASS SPECTROMETRY: MW=1246.56; METHOD=MALDI.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR PROSITE: PS00473; GNRH.
 KW Hormone; Amidation.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
 SQ SEQUENCE 10 AA; 1264 MW; 284B3639DB5AB5A3 CRC64;

Query Match 11.7%; Score 32; DB 1; Length 10;
 Best Local Similarity 55.6%; Pred. No. 2.5e+02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
 ||| : | |
 Db 2 HWSDFKPG 10

RESULT 15

ID UR2_SCYCA STANDARD; PRT; 12 AA.
 AC P35490;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Urotensin II (U-II) (UII).
 OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
 OC Scyllorhinidae; Scyllorhinus.
 OX NCBI_TaxID=7830;

RN [1]
 RP SEQUENCE.
 RC TISSUE=Spinal cord;
 RX MEDLINE=92319231; PubMed=1620290;
 RA Conlon J.M., O'Harte F., Smith D.D., Balment R.J., Hazon N.;
 RT "Purification and characterization of urotensin II and parvalbumin
 from an elasmobranch fish, Scyllorhinus canicula (common dogfish).";
 RL Neuroendocrinology 55:230-235(1992).
 CC -1- FUNCTION: HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
 CC CORTICOTROPIN-RELEASING FACTOR. PROBABLY INVOLVED IN SMOOTH
 CC MUSCLE STIMULATION.
 CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.
 DR InterPro: IPR001483; Urotensin_II.
 DR Pfam: PF02083; Urotensin_II; 1.
 DR PROSITE: PS00984; UROTENSIN_II; 1.
 KW Hormone.
 FT DISULFID 6 11
 SQ SEQUENCE 12 AA; 1526 MW; 804729F9D579CEBA CRC64;

Query Match 11.7%; Score 32; DB 1; Length 12;
 Best Local Similarity 62.5%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 18 NNFTVSEW 25
 ||| : | |
 Db 1 NNFSDCF 8

Search completed: October 10, 2002, 22:58:51
 Job time : 13 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:54:50 ; Search time 28 Seconds
(without alignments)
308.920 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHSYGLRGSGPSLFNPF.....VSASHLEGSLHWSYGLRPX 50

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 29986

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	16.4	33	13	Q9PT34 oncorhynchu
2	45	16.4	33	13	Q9W7G0 oncorhynchu
3	42.5	15.5	49	2	Q9FB81 helicobacte
4	41.5	15.1	32	2	O05602 pseudomonas
5	40.5	14.8	37	5	Q9BM29 saccoglossu
6	40	14.6	31	1	Q55314 sulfolobus
7	40	14.6	30	5	O44051 trypanosoma
8	39	14.2	33	12	Q9LJ04 tt virus. o
9	38.5	14.1	49	2	Q9FB79 helicobacte
10	38	13.9	20	11	Q9EQX8 mus musculu
11	38	13.9	32	5	Q9U340 caenorhabdi
12	37	13.5	40	9	Q37958 bacterioph
13	37	13.5	45	4	Q96B35 homo sapien
14	36.5	13.3	45	7	P79529 homo sapien
15	36.5	13.3	45	7	P79532 homo sapien
16	36.5	13.3	45	7	P79537 homo sapien

17	36	13.1	21	10	Q41566 triticum ae
18	36	13.1	40	11	Q9DIX8 mus musculu
19	36	13.1	40	11	Q99JK9 mus musculu
20	36	13.1	46	4	Q96J56 homo sapien
21	35.5	13.0	29	6	Q9TRN1 sus scrofa
22	35.5	13.0	50	4	Q9UD35 homo sapien
23	35	12.8	25	5	Q9BM38 caenorhabdi
24	35	12.8	29	10	Q9S8N0 raphanus sa
25	35	12.8	38	4	Q96D55 homo sapien
26	35	12.8	38	16	Q9KM73 vibrio chol
27	35	12.8	41	15	Q69891 human immun
28	35	12.8	41	15	Q69892 human immun
29	35	12.8	44	8	Q32439 hordeum vul
30	35	12.8	45	10	Q49787 arabidopsis
31	35	12.8	49	11	Q9EQ69 mus musculu
32	34.5	12.6	33	12	Q69077 human herpe
33	34.5	12.6	47	12	Q69899 human herpe
34	34	12.4	32	6	Q9N1V7 equus cabal
35	34	12.4	32	11	Q9CYM6 mus musculu
36	34	12.4	33	10	Q9XG99 pisum sativ
37	34	12.4	34	11	Q99KX7 mus musculu
38	34	12.4	42	3	O42780 ustilago ho
39	34	12.4	44	7	P79528 homo sapien
40	34	12.4	44	7	P79535 homo sapien
41	34	12.4	45	5	O44828 caenorhabdi
42	34	12.4	46	2	Q93LU8 helicobacte
43	34	12.4	46	2	Q93LU1 helicobacte
44	34	12.4	46	2	Q93LT7 helicobacte
45	34	12.4	46	2	Q93LS2 helicobacte

ALIGNMENTS

RESULT 1

Q9PT34 PRELIMINARY; PRT; 33 AA.

AC Q9PT34; TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
GN GNRH1
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalbun K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalbun K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110533; AAD3461.1; -.
DR InterPro; IPR002047; AKH.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00256; AKH; UNKNOWN_1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;


```
OC Saccorhossus
OX NCBI_TaxID=10224;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-LRT-L3 RETROTRANSPOSON;
MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and ancient asexual taxa.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY013968; AAG5949.1; -
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 37
FT NON_TER 37
SQ SEQUENCE 37 AA; 4116 MW; 7C1F65BCBAFF48C8 CRC64;

Query Match 14.8%; Score 40.5; DB 5; Length 37;
Best Local Similarity 38.7%; Pred. No. 2.2e+02;
Matches 12; Conservative 5; Mismatches 11; Indels 3; Gaps 2;

QY 14 PSENNFT--VSFWLRVPKVSASHLEGPSLH 42
Db 5 PKLFALYVDELSYKLRUAKVGC-HLAGVPIH 34

RESULT 6
Q55314 PRELIMINARY; PRT; 31 AA.
AC Q55314
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE URF2 PROTEIN (FRAGMENT).
GN URF2.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96095144; PubMed=8521845;
RA Jones C.E., Fleming T.M., Cowan D.A., Littlechild J.A., Piper P.W.;
RT "The phosphoglycerate kinase and glyceraldehyde-3-phosphate
dehydrogenase genes from the thermophilic archaeon Sulfolobus
solfataricus overlap by 8bp. Isolation, sequencing of the genes and
expression on Escherichia coli.";
RL Eur. J. Biochem. 233:800-808(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94082761; PubMed=8259927;
RA Arcari P., Russo A.D., Ianniciello G., Gallo M., Bocchini V.;
RT "Nucleotide sequence and molecular evolution of the gene coding for
glyceraldehyde-3-phosphate dehydrogenase in the thermoacidophilic
archaeobacterium Sulfolobus solfataricus.";
RL Biochem. Genet. 31:241-251(1993).
DR EMBL; X80178; CAA56461.1; -
FT NON_TER 31
FT NON_TER 31
SQ SEQUENCE 31 AA; 3554 MW; 9A2538F911C7309A CRC64;

Query Match 14.6%; Score 40; DB 1; Length 31;
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 12 SGPSLFNNFTVSFWLR 27
Db 2 SEPSYF-----VSFWLR 13

RESULT 7
O44051 PRELIMINARY; PRT; 50 AA.
ID O44051
AC O44051
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
```

```
DE SIALIDASE HOMOLOG PRECURSOR (FRAGMENT).
OS Trypanosoma rangelli.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5698;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LDG;
MEDLINE=98115452; PubMed=9455917;
RA Buschiazzo A., Campetella O., Frasch A.C.C.;
RT "Trypanosoma rangelli sialidase: cloning, expression and similarity to
T. cruzi trans-sialidase.";
RL Glycobiology 7:1167-1173(1997).
DR EMBL; U83241; AAC38898.1; -
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 >50 SIALIDASE HOMOLOG.
FT NON_TER 50
FT NON_TER 50
SQ SEQUENCE 50 AA; 5627 MW; C88BDB27CECB484 CRC64;

Query Match 14.6%; Score 40; DB 5; Length 50;
Best Local Similarity 47.8%; Pred. No. 3.5e+02;
Matches 11; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 4 SYGLRPGSSGPSLF--NNFTVSF 24
Db 22 AHALAPGSSRVLEFKRQNSTVVF 44

RESULT 8
Q91J04 PRELIMINARY; PRT; 33 AA.
ID Q91J04
AC Q91J04
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF2 (FRAGMENT).
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=6887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BBY-7;
RA Lin H.-H., Kao J.-H., Lee P.-I., Chen D.-S.;
RT "Early acquisition of TT virus in infants: Probable minor role of
maternal transmission.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF380386; AAK59300.1; -
FT NON_TER 1
FT NON_TER 33
FT NON_TER 33
SQ SEQUENCE 33 AA; 3712 MW; 36BD62745C5D3037 CRC64;

Query Match 14.2%; Score 39; DB 12; Length 33;
Best Local Similarity 37.0%; Pred. No. 3e+02;
Matches 10; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 23 SFWLRVPKVSASHLEGPSLHWSYGLRP 49
Db 3 SEWLSFPRPSAARTPRRGFHASRGRVP 29

RESULT 9
Q9FB79 PRELIMINARY; PRT; 49 AA.
ID Q9FB79
AC Q9FB79;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE OUTER MEMBRANE PROTEIN (FRAGMENT).
GN OIPA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GI2895;
RX MEDLINE=20319061; PubMed=10852959;
RA Yamaoka Y., Kwon D.H., Graham D.Y.;
RT "A W(r) 34,000 proinflammatory outer membrane protein (oipa) of
RL Helicobacter pylori.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:7533-7538(2000).
DR EMBL; AF233675; AAG00394.1; -.
FT NON_TER 1
FT 49
FT SEQUENCE 49 AA; 5272 MW; D733BCB35467B1E7 CRC64;
SQ
Query Match 14.1%; Score 38.5; DB 2; Length 49;
Best Local Similarity 25.0%; Pred. No. 5.5e+02;
Matches 8; Conservative 7; Mismatches 8; Indels 9; Gaps 1;
Qy 16 LFNNFTVSFWLRVPK-----VSASHLEG 38
| : : : : :
Db 5 LTNSLSLFWLHAERNGFYLGILNFAEGSYIQG 36

RESULT 10
Q9EQX8 PRELIMINARY; PRT; 20 AA.
AC Q9EQX8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GLUTAMATE DECARBOXYLASE 65 (FRAGMENT).
OS GAD65.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Makino K., Kobayashi T., Kobayashi T., Sinkawa H., Sakagami H.,
RA Kondo H., Tashiro F., Miyazaki J., Obata K., Tamura S., Yanagawa Y.;
RT "Structure of the Mouse Glutamate Decarboxylase 65 Gene and its
RT Promoter.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032757; BAB20415.1; -.
FT NON_TER 20
FT SEQUENCE 20 AA; 1989 MW; 28F544B0E4C3768D CRC64;
SQ
Query Match 13.9%; Score 38; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 3 WSYGLRPGSSGP 14
| : : | : |
Db 9 WSGSGEDGSADP 20

RESULT 11
Q9U340 PRELIMINARY; PRT; 32 AA.
AC Q9U340;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2001 (TrEMBLrel. 16, Last annotation update)
DE W06G6.9 PROTEIN.
GN W06G6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
```

```
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z83129; CAB63325.1; -.
SQ SEQUENCE 32 AA; 3775 MW; E49D031288C6806A CRC64;
Query Match 13.9%; Score 38; DB 5; Length 32;
Best Local Similarity 41.7%; Pred. No. 3.9e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 2 HWSYGLRPGSSG 13
| : : | : |
Db 18 HWQWAARPGEWG 29

RESULT 12
Q37958 PRELIMINARY; PRT; 40 AA.
ID Q37958;
AC Q37958;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE STRUCTURAL PROTEIN.
OS Bacteriophage PM2.
OC Viruses; dsDNA viruses, no RNA stage; Corticoviridae; Corticovirus.
OX NCBI_TaxID=10661;
RN [1]
RP SEQUENCE FROM N.A.
RA MILLER F.D., Winkfein R.J., Rattner J.B., de Sande J.H.;
RT "Sequence analysis of a PM2-DNA anti-Z-IgG-binding region.";
RL Biosci. Rep. 4:885-895(1984).
RN [2]
RP SEQUENCE FROM N.A.
RA Kivela H.M., Mannisto R.H., Kalkkinen N., Bamford D.H.;
RT "Purification and protein composition of PM2, the first lipid-
RT containing bacterial virus to be isolated.";
RL Virology 0:0-0(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Mannisto R.H., Kivela H.M., Paulin L., Bamford D.H., Bamford J.K.H.;
RT "The complete genome sequence of PM2, the first lipid-containing
RT bacterial virus to be isolated.";
RL Virology 0:0-0(1999).
RN [4]
RP SEQUENCE FROM N.A.
RA Mannisto R.H., Kivela H.M., Paulin L., Bamford D.H., Bamford J.K.H.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; M26134; AAA32194.1; -.
DR EMBL; AF155037; AAD43551.1; -.
KW Structural protein.
SQ SEQUENCE 40 AA; 4389 MW; B138B9941B153B9 CRC64;
Query Match 13.5%; Score 37; DB 9; Length 40;
Best Local Similarity 28.1%; Pred. No. 6.8e+02;
Matches 9; Conservative 8; Mismatches 13; Indels 2; Gaps 1;
Qy 2 HWSYGLRPGSGPSLFNNFTVSFWLRVPKVA 33
| : : | : | | : | : |
Db 11 YFAYGVAISAAGAILAE--YVRDWMRKPKAKS 40

RESULT 13
Q96B35 PRELIMINARY; PRT; 45 AA.
ID Q96B35;
AC Q96B35;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE HYPOTHETICAL 4.9 KDA PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016044; AAH16044.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 45 AA; 4872 MW; D291EDDFLDCBFD86 CRC64;

Query Match 13.3%; Score 37; DB 4; Length 45;
Best Local Similarity 50.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 35 HLEGPSLHWSYG 46
DB 34 HVGPSHOWAAG 45

RESULT 14
P79529 ID P79529 PRELIMINARY; PRT; 45 AA.
AC P79529;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE MHC CLASS I-LIKE MOLECULE PERB11.1-35.1 (FRAGMENT).
OS PERB11.1.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95012455; PubMed=7927538;
RA Leelayuwat C., Townend D.C., Degli-Esposti M.A., Abraham L.J.,
RA Dawkins R.L.;
RT "A new polymorphic and multicopy MHC gene family related to
nonmammalian class I.";
RL Immunogenetics 40:339-351(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97152410; PubMed=8995188;
RA Gaudieri S., Leelayuwat C., Townend D.C., Mullberg J., Cosman D.,
RA Dawkins R.L.;
RT "Allelic and interlocus comparison of the PERB11 multigene family in
the MHC.";
RL Immunogenetics 45:209-216(1997).
DR EMBL; U69966; AAB51790.1; -.
FT NON_TER 1
FT NON_TER 45
SQ SEQUENCE 45 AA; 4861 MW; C904EC188602067C CRC64;

Query Match 13.3%; Score 36.5; DB 7; Length 45;
Best Local Similarity 25.6%; Pred. No. 9.1e+02;
Matches 10; Conservative 10; Mismatches 16; Indels 3; Gaps 2;

QY 2 HW-SYGLRPGSSGSLFNFTVSFWLRVPKVSASHLEGP 39
DB 9 HMOTFHVSAAAAAIF--VIIIFYVRCKCKKTSAAEGP 45

RESULT 15
P79532 ID P79532 PRELIMINARY; PRT; 45 AA.
AC P79532;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
```

```
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE MHC CLASS I-LIKE MOLECULE PERB11.1-46.1 (FRAGMENT).
GN PERB11.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95012455; PubMed=7927538;
RA Leelayuwat C., Townend D.C., Degli-Esposti M.A., Abraham L.J.,
RA Dawkins R.L.;
RT "A new polymorphic and multicopy MHC gene family related to
nonmammalian class I.";
RL Immunogenetics 40:339-351(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97152410; PubMed=8995188;
RA Gaudieri S., Leelayuwat C., Townend D.C., Mullberg J., Cosman D.,
RA Dawkins R.L.;
RT "Allelic and interlocus comparison of the PERB11 multigene family in
the MHC.";
RL Immunogenetics 45:209-216(1997).
DR EMBL; U69966; AAB51793.1; -.
FT NON_TER 1
FT NON_TER 45
SQ SEQUENCE 45 AA; 4861 MW; C904EC188602067C CRC64;

Query Match 13.3%; Score 36.5; DB 7; Length 45;
Best Local Similarity 25.6%; Pred. No. 9.1e+02;
Matches 10; Conservative 10; Mismatches 16; Indels 3; Gaps 2;

QY 2 HW-SYGLRPGSSGSLFNFTVSFWLRVPKVSASHLEGP 39
DB 9 HMOTFHVSAAAAAIF--VIIIFYVRCKCKKTSAAEGP 45

Search completed: October 10, 2002, 22:59:41
Job time : 29 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:53:25 ; Search time 28 Seconds
(without alignments)
198.346 Million cell updates/sec

Title: US-09-848-834A-18
Perfect score: 274
Sequence: 1 XHWSYGLRPGSSGSLFNNE.....VSASHLEGPSLHWSYGLRPX 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 352077

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	152	55.5	32	15 AAR62702
2	118.5	43.2	31	21 AAY92655
3	116	42.3	31	21 AAY92653
4	116	42.3	43	22 AAB46177
5	116	42.3	43	22 AAB49076
6	113	41.2	36	22 AAG63662
7	113	41.2	36	22 AAG63515
8	112	40.9	21	12 AAR11896
9	112	40.9	21	17 AAW06130
10	112	40.9	21	17 AAR88397
11	112	40.9	21	18 AAW46449

12	112	40.9	21	19 AAW67034	Tetanus toxin frag
13	112	40.9	21	20 AAW67579	T-cell epitope pep
14	112	40.9	21	20 AAW73222	Tetanus toxoid epi
15	112	40.9	21	21 AAB45512	Tetanus P30 epitop
16	112	40.9	21	21 AAY99876	Tetanus toxin T ce
17	112	40.9	21	21 AAY92626	Foreign epitope P3
18	112	40.9	21	21 AAY84428	Amino acid sequenc
19	112	40.9	21	21 AAY49260	CD4+ T cell epitop
20	112	40.9	21	22 AAE11764	Clostridium tetani
21	112	40.9	21	22 AAB85702	Amino acid sequenc
22	112	40.9	21	22 AAB85453	Universal tetanus
23	112	40.9	21	22 AAB61958	Tetanus toxoid uni
24	112	40.9	21	22 AAB20144	Tetanus toxin T-ce
25	112	40.9	21	22 AAB68637	HER-2 B cell pepti
26	112	40.9	21	22 AAB46173	Tetanus toxoid TT9
27	112	40.9	21	22 AAB49072	Tetanus toxoid TT
28	112	40.9	28	22 AAB46176	Tetanus toxoid 947
29	112	40.9	31	21 AAY92654	PSMpep011 - P30 in
30	112	40.9	33	22 AAB49075	Amyloid beta/tetan
31	112	40.9	44	22 AAB46194	Tetanus toxoid epi
32	112	40.9	44	22 AAB49090	Amyloid beta/tetan
33	111	40.5	34	16 AAR83562	IgE CH4 region con
34	108	39.4	32	16 AAR82597	IgE CH4 region con
35	107	39.1	22	15 AAR62693	Helper T cell epit
36	107	39.1	22	16 AAR82574	Tetanus toxin help
37	107	39.1	22	17 AAW05600	Tetanus toxin help
38	107	39.1	22	21 AAY80057	Pathogen derived T
39	107	39.1	22	21 AAY54540	T helper cell (Th)
40	107	39.1	22	21 AAY58769	Unidentified pepti
41	107	39.1	22	22 AAB84436	Amino acid sequenc
42	106	38.7	21	21 AAY96458	Tetanus toxin (TTD
43	103	37.6	19	21 AAY99061	HLA class II bindi
44	101	36.9	49	17 AAW03944	GNRH 4-repeat sequ
45	101	36.9	49	19 AAW79567	GNRH-2. Synthetic

ALIGNMENTS

RESULT 1	
AAR62702	ID AAR62702 standard; peptide; 32 AA.
XX	AC
XX	AAR62702;
DT	10-SEP-1995 (first entry)
XX	LHRH-containing immunogenic peptide.
DE	Helper T cell epitope; universal immune stimulator; invasive; hapten.
XX	KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW	androgen-dependent carcinoma; antitumour; infertility;
KW	tetanus toxin.
XX	Synthetic.
XX	Key
FT	Domain
FT	Domain
FT	Domain
FT	Domain
FT	Domain
XX	Location/Qualifiers
XX	1..22 "tetanus toxin helper T cell epitope"
XX	23..32
XX	/note= "LHRH haptent"
XX	WO9425060-A.
XX	10-NOV-1994.
XX	28-APR-1994; 94WO-US04832.
XX	27-APR-1993; 93US-0057166.
XX	14-APR-1994; 94US-0229275.
XX	(LADD/) LADD A E.
XX	(WANG/) WANG C Y.

```

PA (ZAMB/) ZAMB T.
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX Immunogenic luteinising hormone releasing hormone peptide(s) -
PT that suppress LHRH activity in males and females
XX Claim 8; Page 84; 213pp; English.
XX Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein haptan containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The
CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
CC which elicits an immune response to the coupled peptide in members of
CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
CC an adjuvant peptide sequence from the invasive protein of Yersinia.
CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
CC invasin and Th domains and between the immune stimulator and haptan
CC components. When the haptan is LHRH, then optionally the invasin domain
CC can be omitted from the immune stimulator component.
CC The present sequence represents an LHRH-containing, invasin-free
CC immunogenic peptide as above which can be used as a potent vaccine for
CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
CC uterine tumours, recurrent functional ovarian cysts, (severe)
CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
CC induction of infertility.
XX SQ Sequence 32 AA;
Query Match 55.5%; Score 152; DB 15; Length 32;
Best Local Similarity 87.9%; Pred. No. 3.4e-13;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
QY 17 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 49
DB 3 FNNFTVFWLRVPKVSASHLE---HWSYGLRP 31
RESULT 2
AA92655
ID AAY92655 standard; Peptide; 31 AA.
XX AC AAY92655;
XX DT 10-AUG-2000 (first entry)
XX DE PSMpep012 - P30 inserted in hPSM insertion position 10.
XX KW Foreign epitope; P2; prostate specific membrane antigen; vaccination;
KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW prostate cancer; cell-associated peptide antigen.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 6..26
XX FT /label= P30
XX PN WO200020027-A2.
XX PD 13-APR-2000.
XX PF 05-OCT-1999; 99WO-DK00525.
XX PR 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX PR (MEBI-) M & E BIOTECH AS.
XX PA

XX Steinna L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX WPI; 2000-349917/30.
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX Example 1; Page 118; 220pp; English.
XX AAY92650-55 are peptides designed which correspond to the P2 and P30
CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
CC amino acids in each end. The flanking amino acids correspond to the
CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g.
CC T cell proliferation assays, but also for ELISA or other in vitro
CC assays. The claims detail a method for inducing immune responses against
CC weakly immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, hPSM,
CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
CC method comprises effecting simultaneous presentation by antigen
CC producing cells (APCs) of the animals immune system of: (1) at least 1
CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC B-cell group derived from the cell-associated PA; and (2) at least 1
CC first T helper cell group which is foreign to the animal. Analogues of
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC part of all known and predicted CTL and B-cell epitopes of the
CC respective PA and including at least one foreign T helper epitope (e.g.
CC P2 and/or P30) are also claimed. The method is used to treat prostate,
CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and
CC Her2, respectively.
XX SQ Sequence 31 AA;
Query Match 43.2%; Score 118.5; DB 21; Length 31;
Best Local Similarity 92.3%; Pred. No. 9.4e-09;
Matches 24; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 17 FNNFTVFWLRVPKVSASHLEGPSLH 42
DB 6 FNNFTVFWLRVPKVSASHLE-PSH 30
RESULT 3
AA92653
ID AAY92653 standard; Peptide; 31 AA.
XX AC AAY92653;
XX DT 10-AUG-2000 (first entry)
XX DE PSMpep010 - P30 inserted in hPSM insertion position 6.
XX KW Foreign epitope; P2; prostate specific membrane antigen; vaccination;
KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW prostate cancer; cell-associated peptide antigen.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 6..26
XX FT /label= P30
XX PN WO200020027-A2.
XX PD 13-APR-2000.
XX PF 05-OCT-1999; 99WO-DK00525.
XX PR 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX PR
XX

```

PA (MEBI-) M & E BIOTECH AS.

XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX WPI; 2000-349917/30.
DR

XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
PS Example 1; Page 117; 220pp; English.

XX AAY92650-55 are peptides designed which correspond to the P2 and P30
CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
CC amino acids in each end. The flanking amino acids correspond to the
CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g.
CC T cell proliferation assays, but also for ELISA or other in vitro
CC assays. The claims detail a method for inducing immune responses against
CC weakly immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, hPSM,
CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
CC method comprises effecting simultaneous presentation by antigen
CC producing cells (APCs) of the animals immune system of: (1) at least 1
CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC B-cell group derived from the cell-associated PA; and (2) at least 1
CC first T helper cell group which is foreign to the animal. Analogues of
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC part of all known and predicted CTL and B-cell epitopes of the
CC respective PA and including at least one foreign T helper epitope (e.g.
CC P2 and/or P30) are also claimed. The method is used to treat prostate,
CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and
CC Her2, respectively.
XX

SQ Sequence 31 AA;

Query Match 42.3%; Score 116; DB 21; Length 31;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LFNNFTVSFWLRVPKVSASHLE 37
DB 5 LFNNFTVSFWLRVPKVSASHLE 26
|||||

RESULT 4
AAB46177

ID AAB46177 standard; peptide; 43 AA.

XX AAB46177;

XX 04-APR-2001 (first entry)

DE Tetanus toxoid 830-844 + 947-967 epitope AN90542.

XX Amyloid deposit; APP; Abeta; brain; human; clearing response; neurotropic;
KW Fc receptor mediated phagocytosis; immunogenic response; neuroprotective;
KW amyloid precursor protein; Alzheimer's disease.
XX

OS Clostridium tetani.

XX WO200072880-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14810.

XX 28-MAY-1999; 99US-0322289.

XX (NEUR-) NEURALAB LTD.

XX Schenk DB, Bard F, Vasquez NJ, Yednock T;

XX WPI; 2001-032104/04.

XX

PT Preventing or treating a disease associated with amyloid deposits,
PT especially Alzheimer's disease, comprises administering amyloid
PT specific antibody -

XX Disclosure; Page 31; 143pp; English.

XX

CC This invention describes a novel method of preventing or treating a
CC disease associated with amyloid deposits of amyloid precursor protein
CC (APP) Abeta fragments in the brain of a patient, which comprises
CC administering to the patient: (a) an antibody that binds to Abeta, the
CC antibody binds to an amyloid deposit and induces a clearing response (Fc
CC receptor mediated phagocytosis) against it (b) a polypeptide containing
CC an N-terminal segment of at least residues 1-5 of Abeta; or (c) an agent
CC that induces an immunogenic response against residues 1-3 to 7-11 of
CC Abeta. The products of the invention have neurotropic and neuroprotective
CC activity. The method is also useful for monitoring a course of treatment
CC being administered to a patient e.g. active and passive immunization. The
CC methods are useful for prophylactic and therapeutic treatment of
CC Alzheimer's disease.

XX Sequence 43 AA;

Query Match 42.3%; Score 116; DB 22; Length 43;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LFNNFTVSFWLRVPKVSASHLE 37
DB 22 LFNNFTVSFWLRVPKVSASHLE 43
|||||

RESULT 5
AAB49076

ID AAB49076 standard; peptide; 43 AA.

XX AAB49076;

XX 27-MAR-2001 (first entry)

DE Amyloid beta/tetanus toxoid immunogenic fusion peptide, SEQ ID NO:12.

XX Amyloid disease; amyloid fibril deposition; amyloid plaque;
KW immunogenic; antibody; vaccine; Alzheimer's disease;
KW type 2 diabetes; reactive system amyloidosis;
KW systemic senile amyloidosis; familial amyloid cardiomyopathy;
KW transmissible spongiform encephalopathy; Creutzfeldt-Jakob disease; Kuru;
KW haemodialysis-associated beta-2-microglobulin deposition;
KW amyloid beta peptide; universal T-cell epitope; neuroprotective.

XX Chimeric - Homo sapiens.

OS Chimeric - Clostridium tetani.

XX WO200072876-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US15239.

XX 01-JUN-1999; 99US-0137010.

XX (NEUR-) NEURALAB LTD.

XX Schenk DB;

XX WPI; 2001-070921/08.

XX Pharmaceutical composition comprising immunogen against amyloid
PT component such as fibril peptide or protein, or antibody against
PT amyloid component useful for treating amyloid diseases or amyloidoses -
XX Disclosure; Page 45; 140pp; English.

CC The invention relates to a novel pharmaceutical composition for
 CC preventing or treating a disease characterised by amyloid fibril
 CC deposits (amyloid plaques) in a patient. The pharmaceutical composition
 CC comprises an agent that will induce an immune response against an amyloid
 CC component, or an antibody or antibody fragment that binds to an amyloid
 CC component. The invention also relates to a method for determining
 CC the prognosis of a patient undergoing treatment for an amyloid disorder
 CC which involves measuring a patient serum amount of immunoreactivity
 CC against a selected amyloid component. A patient serum immunoreactivity
 CC of at least four times a base line serum immunoreactivity control level
 CC indicates a prognosis of improved status with respect to the disorder.
 CC The pharmaceutical compositions of the invention are useful for treating
 CC a wide variety of disorders characterised by amyloid fibril deposition in
 CC a patient. Such disorders include Alzheimer's disease characterised by
 CC amyloid beta peptide fibril deposits; type 2 diabetes characterised by
 CC islet amyloid protein peptide (IAPP, amylin) fibrils; reactive systemic
 CC amyloidosis associated with systemic inflammatory diseases (e.g.,
 CC rheumatoid arthritis, osteomyelitis, tuberculosis) characterised by AA
 CC fibrils derived from serum amyloid A protein (ApoSAA); systemic senile
 CC amyloidosis and familial amyloid cardiomyopathy characterised by ATTR
 CC fibrils derived from transthyretin (TTR); transmissible spongiform
 CC encephalopathies (e.g. Creutzfeldt-Jakob disease, Kuru) characterised by
 CC prion protein deposits; and beta-2-microglobulin deposits which form as
 CC a result of long term haemodialysis treatment. The present sequence
 CC represents an immunogenic fusion protein comprising an amyloid beta
 CC peptide fused to a universal T-cell epitope which may be used in a
 CC composition to treat or prevent Alzheimer's disease.

XX Sequence 43 AA;

Query Match 42.3%; Score 116; DB 22; Length 43;

Best Local Similarity 100.0%; Pred. No. 3e-08;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVFWLRPKVSASHLE 37

DB 22 LFNNFTVFWLRPKVSASHLE 43

RESULT 6

AAG63662

ID AAG63662 standard; peptide; 36 AA.

XX

AC AAG63662;

XX

DT 29-OCT-2001 (first entry)

XX

DE Peptide comprising 5 conjugation sites for a pseudo-peptide.

XX

KW Pseudo-peptide; immunomodulation; adjuvant; antigen presenting cell;

XX

KW macrophage; dendritic cell; vaccine; autoimmune disease.

XX

OS Synthetic.

XX

PN WO200146127-A1.

XX

PD 28-JUN-2001.

XX

PF 22-DEC-1999; 99WO-IB02038.

XX

PR 22-DEC-1999; 99WO-IB02038.

XX

PA (OMPH-) OM-PHARMA.

XX

PI Bauer J, Martin OR, Rodriguez S;

XX

DR WPI; 2001-502469/55.

XX

PT New amphiphilic acylated pseudo-peptides having a functionalized

XX

PT vaccines

XX

PS Example 3; Page 61; 166pp; French.

XX

CC The specification describes N-Acylated pseudo-peptides, which have
 CC a neutral or charged acidic group at one terminal and a functionalized
 CC auxiliary spacer at the other. The pseudo-peptides show immunomodulatory
 CC and adjuvant action, based on activation of antigen presenting cells
 CC (e.g. macrophages or dendritic cells), induction of differentiation of
 CC dendritic cells, induction of cytokine production and induction of
 CC maturation of immunocompetent cell strains originating from hematopoietic
 CC and lymphoid organs. They reinforce humoral and cellular immunity. They
 CC can be grafted onto antigens (to modulate immune response) or onto drugs
 CC (to improve the therapeutic activity or targeting). The pseudo-peptides
 CC are thus useful in human or veterinary medicine as immunizing or
 CC diagnostic agents. Typically, they are used as adjuvants together with
 CC (or covalently bonded to) antigens for vaccination against viral,
 CC parasitic/protozoal, microbial or fungal infections; incubated with blood
 CC cells ex vivo, to render the cells immunocompetent before reintroduction
 CC in vivo; or used in therapy of certain autoimmune diseases. The
 CC pseudo-peptides are useful as carriers for antigens or other therapeutic
 CC agents due to their ability to form non-covalent bonds via the
 CC hydrophobic or hydrophilic auxiliary spacer. The present sequence
 CC represents a peptide, which has 5 possible conjugation sites for the
 CC pseudo-peptides of the invention.

XX Sequence 36 AA;

Query Match 41.2%; Score 113; DB 22; Length 36;

Best Local Similarity 88.0%; Pred. No. 6e-08;

Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 GPSLFNNFTVFWLRPKVSASHLE 37

DB 12 GITEFNNFTVFWLRPKVSASHLE 36

RESULT 7

AAG63515

ID AAG63515 standard; peptide; 36 AA.

XX

AC AAG63515;

XX

DT 15-OCT-2001 (first entry)

XX

DE A peptide which may be conjugated to pseudo-peptides.

XX

KW Pseudo-peptide; immunomodulation; adjuvant; antigen presenting cell;

XX

KW macrophage; dendritic cell; cytokine production; immunocompetent cell;

XX

OS Synthetic.

XX

PN WO200146126-A1.

XX

PD 28-JUN-2001.

XX

PF 21-DEC-2000; 2000WO-FR03650.

XX

PR 22-DEC-1999; 99WO-IB02038.

XX

PA (OMPH-) OM-PHARMA.

XX

PI Bauer J, Martin OR, Rodriguez S;

XX

DR WPI; 2001-496651/54.

XX

PT New amphiphilic acylated pseudo-peptides having a functionalized

XX

PT auxiliary spacer, useful as immunomodulators, e.g. as adjuvants in

XX

PS Example 3.3; Page 88; 267pp; French.

XX

CC The specification describes N-Acylated pseudo-peptides, which have
 CC a neutral or charged acidic group at one terminal and a functionalized
 CC auxiliary spacer at the other. The pseudo-peptides show immunomodulatory

CC and adjuvant action, based on activation of antigen presenting cells
 CC (e.g. macrophages or dendritic cells), induction of differentiation of
 CC dendritic cells, induction of cytokine production and induction of
 CC maturation of immunocompetent cell strains originating from hematopoietic
 CC and lymphoid organs. They reinforce humoral and cellular immunity. They
 CC can be grafted onto antigens (to modulate immune response) or onto
 CC drugs (to improve the therapeutic activity or targeting). The
 CC pseudopeptides are thus useful in human or veterinary medicine as
 CC immunizing or diagnostic agents. Typically, the pseudopeptides are used
 CC as adjuvants together with (or covalently bonded to) antigens for
 CC vaccination against viral, parasitic/protozoal, microbial or fungal
 CC infections; incubated with blood cells ex vivo; or used in therapy of
 CC immunocompetent before reintroduction in vivo; or used in therapy of
 CC certain autoimmune diseases. The present sequence represents a
 CC peptide which may be conjugated to pseudopeptides of the invention.

SQ Sequence 36 AA;

Query Match 41.2%; Score 113; DB 22; Length 36;
 Best Local Similarity 88.0%; Pred. No. 6e-08;
 Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 GPSLENNFTVSWLRVPKVSASHLE 37
 | : ||||| ||||| ||||| |||||
 DB 12 GITEFNFTVSWLRVPKVSASHLE 36

RESULT 8

AAR11896
 ID AAR11896 standard; peptide; 21 AA.

XX AAR11896;

DT 19-JUL-1991 (first entry)

XX Immunogenic conjugate constituent peptide, TT3.

XX Malaria vaccine; major histocompatibility complex.

XX Tetanus toxin.

FH Key Location/Qualifiers
 FT Peptide 1..14
 /label= active fragment (claimed)

XX EP427347-A.

XX 15-MAY-1991.

XX 07-NOV-1990; 90EP-0202948.

XX 10-NOV-1989; 89IT-0022355.

XX (ENTE) ENTRICERCH SPA.

XX Bianchi E, Pessi A, Corradin G;

XX WPI; 1991-141874/20.

XX Synthetic peptide(s) used as universal carriers - for preparing
 PT immunogenic conjugates used as vaccines against Plasmodium
 PT falciparum

XX Claim 1; page 13; 16pp; English.

XX This peptide corresponds to residues 947-967 of Tetanus toxin. It can
 CC be used as a universal carrier for the prepn. of an immunogenic
 CC conjugate. It is covalently bound to a peptide or polysaccharide
 CC hapten derived from a pathogen. This conjugate can be used as
 CC a vaccine for malaria. This peptide is recognised by different T-
 CC helper cell clones in association with alleles of the human MHC.
 CC It contains 2 epitopes: (a) 953-967, recognised by DR5-restricted
 CC clones; and (b) 947-960, recognised by all other DR and Dp-

CC restricted clones.
 XX SQ Sequence 21 AA;

Query Match 40.9%; Score 112; DB 12; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRVPKVSASHLE 37
 ||||| ||||| ||||| |||||
 DB 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 9

AAW06130
 ID AAW06130 standard; Peptide; 21 AA.

XX AAW06130;

DT 07-FEB-1997 (first entry)

XX Tetanus toxoid protein T-cell epitope.

XX Cholesteryl ester transfer protein; CETP; antigen; vaccine;

KW cardiovascular disease; atherosclerosis; tetanus toxoid;

KW T-cell epitope.

XX Clostridium tetani.

XX WO9634888-Al.

XX 07-NOV-1996.

XX 01-MAY-1996; 96WO-US06147.

XX 01-MAY-1995; 95US-0432483.

XX (TCEL-) T CELL SCI INC.

XX Rittershaus CW, Thomas LJ;

XX WPI; 1996-506103/50.

XX Cholesteryl ester transfer protein B cell epitope linked to T cell
 PT epitope - used to generate vaccine to regulate CETP activity for
 PT decreasing the risk of developing a cardiovascular disease e.g.
 PT atherosclerosis

XX Claim 11; Page 43; 72pp; English.

XX A helper T-cell epitope (AAW06130) comprises amino acids 947-967
 CC of tetanus toxoid protein. It can be utilised in novel peptide
 CC vaccines (see also AAW06129, AAW06132) also including B-cell
 CC epitope(s) from human or rabbit cholesteryl ester transfer
 CC protein (CETP) to elicit an immune response against endogenous
 CC CETP activity, thereby treating or preventing a cardiovascular
 CC disease, such as atherosclerosis.

XX SQ Sequence 21 AA;

Query Match 40.9%; Score 112; DB 17; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRVPKVSASHLE 37
 ||||| ||||| ||||| |||||
 DB 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 10

AAAR88397
 ID AAR88397 standard; Peptide; 21 AA.

XX

```
AC AAR88397;
XX
XX 12-JUN-1996 (first entry)
XX
XX T-cell antigen TT3 peptide.
XX
XX T-antigen; vaccine; antibody; T-cell; T-lymphocyte;
KW alpha-helix; coiled-coil heterodimer; core peptide; subunit.
XX
XX Synthetic.
XX
XX WO9531480-AL.
XX
XX 23-NOV-1995.
XX
XX 18-MAY-1995; 95WO-CA00293.
XX
XX 18-MAY-1994; 94US-0245507.
XX
XX (SPIS-) SPI SYNTHETIC PEPTIDES INC.
XX
XX Cachia PJ, Hodges RS, Houston ME, Irvin RT, Kay CM;
PI Zhou NE;
XX
XX WPI; 1996-010880/01.
XX
XX Hetero:dimeric polypeptide immunogen in coiled-coil configuration
PT with different antigens on each sub:unit - useful in vaccines and
PT for antibody prodn.
XX
XX Claim 7; Page 62; 95pp; English.
XX
XX This T-cell antigen TT3 peptide may be attached to a core peptide
CC contained in one of the 2 subunits of an alpha-helical coiled-coil
CC heterodimer. Each core peptide is comprised of terminal and
CC internal AA repeat sequences. This peptide antigen is attached
CC to the core peptide through covalent linkages to certain AA of the
CC in internal repeats. The 2 subunits of the heterodimer are arranged
CC in a stable alpha-helical coiled-coil configuration having a 1:1
CC stoichiometry, and the peptide antigen is disposed toward the outer
CC surfaces of the configuration. The heterodimer may be used as a
CC synthetic vaccine (optionally multivalent) or to generate
CC antibodies.
XX
XX SQ Sequence 21 AA;
Query Match 40.9%; Score 112; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
RESULT 11
AAW46449
ID AAW46449 standard; Peptide; 21 AA.
XX
XX AAW46449;
XX
XX 18-MAY-1998 (first entry)
XX
XX Broad range helper T cell epitope from the tetanus toxoid protein.
XX
XX Cholesteryl ester transfer protein; CETP; cholesteryl ester;
XX high density lipoprotein; HDL; very low density lipoprotein; VLDL;
XX low density lipoprotein; LDL; T cell epitope; antibody;
XX DNA plasmid-based vaccine; broad range helper T cell epitope;
XX treatment; cardiovascular disease.
XX
XX OS Clostridium tetani.
XX
```

```
PN WO9741227-AL.
XX
XX 06-NOV-1997.
XX
XX 01-MAY-1997; 97WO-US07294.
XX
XX 21-FEB-1997; 97US-0802967.
PR 01-MAY-1996; 96US-0640713.
XX
XX (TCEL-) T CELL SCI INC.
XX
XX Thomas LJ;
PI
XX WPI; 1997-549731/50.
XX
XX DNA plasmid-based vaccine encodes CETP B cell and helper T cell
PT epitope(s) - used for elevating high density lipoprotein levels, and
PT for treating cardiovascular disease
XX
XX Disclosure; Page 44; 67pp; English.
XX
XX The present sequence represents a broad range helper T cell epitope
CC of the tetanus protein. It can be used in DNA plasmid-based vaccines
CC against cholesteryl ester transfer proteins (CETPs). CETPs mediate the
CC transfer of cholesteryl esters from high density lipoprotein (HDL)
CC to very low density lipoprotein (VLDL) and low density lipoprotein (LDL),
CC and vice versa. An increased CETP activity produces an atherogenic
CC lipoprotein profile and induces atherosclerosis. A DNA plasmid-based
CC vaccine comprises sequences encoding at least one B cell epitope of CETP
CC linked in frame with at least one segment encoding a broad range helper
CC T cell epitope. The vaccines can be used to elevate the ratio of
CC circulating HDL to circulating LDL, VLDL or total cholesterol in a human.
CC It can also be used for decreasing the level of endogenous CETP activity
CC in a human. The vaccine can be used to produce anti-CETP antibodies in
CC vivo and for treating cardiovascular disease.
XX
XX SQ Sequence 21 AA;
Query Match 40.9%; Score 112; DB 18; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
RESULT 12
AAW67034
ID AAW67034 standard; peptide; 21 AA.
XX
XX AAW67034;
XX
XX 15-DEC-1998 (first entry)
XX
XX Tetanus toxin fragment (residues 947-967).
XX
XX Tetanus toxin; vaccine; antibody; carbohydrate peptide conjugate;
KW dendrimeric poly-lysine; epitope; tumour.
XX
XX OS Clostridium tetani.
XX
XX WO9843677-AL.
PN
XX 08-OCT-1998.
XX
XX 27-MAR-1998; 98WO-EP01922.
PF
XX 27-MAR-1997; 97US-0041726.
PR
XX (INSP ) INST PASTEUR.
XX
XX Bay S, Cantacuzene D, Leclerc C, Lo-man R;
PI
```

XX WPI; 1998-557071/47.
 XX Carbohydrate peptide conjugate used as vaccine - comprises carrier
 PT with dendrimeric poly-lysine enabling multiple epitopes to be
 PT covalently attached
 XX
 PS Disclosure; Page 13; 55pp; English.
 XX
 CC The invention relates to a new carbohydrate peptide conjugate, which
 CC comprises a carrier with a dendrimeric poly-lysine enabling multiple
 CC epitopes to be covalently attached to it. Also claimed are: (1) an
 CC antibody purified from biological fluid or cells of organisms
 CC administered with the carbohydrate peptide conjugate, and (2) a diagnosis
 CC kit comprising antigen-specific antibodies elicited by immunisation with
 CC the carbohydrate peptide conjugate. The peptide conjugate, antibody and
 CC diagnosis kit are used to provide pharmaceutical compositions and
 CC vaccines against tumours. These can be used to support an immune response
 CC against viral infections caused by hepatitis virus, HIV or cytomegalo
 CC virus. They can be used to enhance immune responses, especially B- and T-
 CC cell responses, of humans and animals against bacterial infections. The
 CC carbohydrate peptide conjugate stimulates the antibody and T-cell
 CC response without stimulating undesired immune responses. The composition
 CC is capable of increasing the survival of tumour bearing humans and
 CC animals. The present sequence corresponds to residues 947-967 of tetanus
 CC toxin. The synthetic peptide corresponding to this sequence may be used
 CC as an epitope in a carbohydrate peptide conjugate.
 XX
 SQ Sequence 21 AA;
 Query Match 40.9%; Score 112; DB 19; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 FNNFTVSEFLRVPKVSASHLE 37
 Db 1 FNNFTVSEFLRVPKVSASHLE 21
 RESULT 13
 AAW67579
 ID AAW67579 standard; peptide; 21 AA.
 XX
 AC AAW67579;
 XX
 DT 02-MAR-1999 (first entry)
 XX
 DE T-cell epitope peptide #5 for chimeric fimbrin/T-cell epitope peptide.
 XX
 KW Chimeric; non-typable Haemophilus influenzae; fimbrin; T-cell epitope;
 KW immunogenic composition; immune response.
 XX
 OS Synthetic.
 XX
 PN US5843464-A.
 XX
 PD 01-DEC-1998.
 XX
 PF 02-JUN-1995; 95US-0460502.
 XX
 PR 02-JUN-1995; 95US-0460502.
 XX
 PA (OHIS) UNIV OHIO STATE.
 XX
 PI Bakaletz LO, Kaumaya PTP;
 XX
 DR WPI; 1999-044514/04.
 XX
 PT Synthetic chimeric fimbrin peptide - useful for vaccination against
 PT non-typable Haemophilus influenzae
 XX
 PS Disclosure; Column 4; 16pp; English.
 XX

CC The invention relates to the manufacture of a synthetic chimeric peptide
 CC comprising a non-typable Haemophilus influenzae fimbrin peptide fused via
 CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is
 CC used in immunogenic compositions which induce an immune response against
 CC non-typable Haemophilus influenzae. This sequence represents an example
 CC of a T-cell epitope peptide used to generate the chimeric peptide.
 XX
 SQ Sequence 21 AA;
 Query Match 40.9%; Score 112; DB 20; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.3e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 FNNFTVSEFLRVPKVSASHLE 37
 Db 1 FNNFTVSEFLRVPKVSASHLE 21
 RESULT 14
 AAW73222
 ID AAW73222 standard; Protein; 21 AA.
 XX
 AC AAW73222;
 XX
 DT 25-JAN-1999 (first entry)
 XX
 DE Tetanus toxoid epitope.
 XX
 KW Multispecific single chain antibody; antibody H22; tumour cell; therapy;
 KW antibody-dependent cellular cytotoxicity; ADCC; HER 2/neu; infection;
 XX epidermal growth factor receptor; breast cancer; ovarian cancer.
 OS Synthetic.
 XX
 PN US5837243-A.
 XX
 PD 17-NOV-1998.
 XX
 PF 07-JUN-1996; 96US-0661052.
 XX
 PR 07-JUN-1996; 96US-0661052.
 PR 07-JUN-1995; 95US-0484172.
 XX
 PA (MEDA-) MEDAREX INC.
 XX
 PI Deo YM, Goldstein J, Graziano R, Somasundaram C;
 XX
 DR WPI; 1999-023374/02.
 XX
 PT Specific killing of tumour cells - using a multi-specific molecule
 PT comprising an anti-Fc receptor antibody and a portion which binds to
 PT a target cell
 XX
 PS Example 7; Column 27; 57pp; English.
 XX
 CC This sequence represents a tetanus toxoid epitope and is recognised
 CC by the multispecific single chain antibody designated H22. The
 CC antibody can be used in the method of the invention for inducing
 CC antibody-dependent cellular cytotoxicity (ADCC) against a tumour cell
 CC which is characterised by overexpression of HER 2/neu or epidermal growth
 CC factor receptor (EGFR), comprises contacting the tumour cell with a
 CC multispecific protein molecule (preferably a single chain antibody)
 CC comprising: (a) an anti-Fc receptor antibody or an antigen binding
 CC fragment; (b) a portion which binds to HER 2/neu; and (c) a portion which
 CC binds to EGFR. The method can be used for treating cancers especially
 CC breast cancer or ovarian cancer. The multispecific antibody can also
 CC be administered prophylactically to vaccinate a subject against infection
 CC by a target cell.
 XX
 SQ Sequence 21 AA;
 Query Match 40.9%; Score 112; DB 20; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.3e-08;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
 |||||
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 15

AAB45512
 ID AAB45512 standard; Protein; 21 AA.

XX AC AAB45512;

XX DT 26-FEB-2001 (first entry)

XX DE Tetanus P30 epitope SEQ ID NO: 24.

XX KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
 KW cancer; eosinophilia; vaccine; allergic rhinitis.

XX OS Clostridium tetani.

XX PN WC200065058-A1.

XX PD 02-NOV-2000.

XX PF 19-APR-2000; 2000WO-DK00205.

XX PR 23-APR-1999; 99DK-0000552.

XX PR 06-MAY-1999; 99US-0132811.

XX PA (MEBI-) M & E BIOTECH AS.

XX PI Klysner S;

XX DR WPI; 2000-672791/65.

XX PT Down-regulating interleukin 5 (IL-5) activity in humans by
 PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
 PT prophylaxis or amelioration of asthma or other chronic allergic
 PT conditions -

XX PS Example 6; Page 137; 172pp; English.

CC The present invention is concerned with methods of treating asthma,
 CC eosinophilia, allergic rhinitis and other allergic diseases. These
 CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
 CC proteins and their coding sequences to down-regulate IL-5 activity and
 CC thus reduce eosinophil numbers. The allergic diseases may be treated
 CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
 CC it is possible that they may be used in the treatment of cancer and
 CC helminthic infections.

XX SQ Sequence 21 AA;

Query Match 40.9%; Score 112; DB 21; Length 21;

Best Local Similarity 100.0%; Pred. No. 4.3e-08;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
 |||||
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

Search completed: October 10, 2002, 22:58:18

Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:57:35 ; Search time 16 Seconds
(without alignments)
76.330 Million cell updates/sec

Title: US-09-848-834A-18
Perfect score: 274
Sequence: 1 XHWSYGLRPGSSGSLFNNF.....VSASHLEGLSLHWSYGLRPX 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 164183

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	55.5	32	1 US-08-446-692-14	Sequence 14, Appl
2	152	55.5	32	2 US-08-488-351A-14	Sequence 14, Appl
3	114	41.6	31	5 PCT-US93-11703-64	Sequence 64, Appl
4	112	40.9	21	1 US-07-610-525-1	Sequence 1, Appl
5	112	40.9	21	2 US-08-661-052-12	Sequence 12, Appl
6	112	40.9	21	2 US-08-460-502-8	Sequence 8, Appl
7	112	40.9	21	2 US-08-724-774B-5	Sequence 5, Appl
8	112	40.9	21	4 US-09-089-595-5	Sequence 5, Appl
9	112	40.9	21	4 US-09-382-855-5	Sequence 5, Appl
10	112	40.9	21	4 US-09-183-714B-5	Sequence 5, Appl
11	112	40.9	21	4 US-09-188-082-12	Sequence 12, Appl
12	112	40.9	21	4 US-09-171-969-10	Sequence 10, Appl
13	112	40.9	21	5 PCT-US93-11703-66	Sequence 66, Appl
14	107	39.1	22	1 US-08-446-692-5	Sequence 5, Appl
15	107	39.1	22	2 US-08-488-351A-5	Sequence 5, Appl
16	107	39.1	22	3 US-09-100-409A-41	Sequence 41, Appl
17	107	39.1	22	5 PCT-US95-13841-8	Sequence 8, Appl
18	101	36.9	49	1 US-08-387-156-4	Sequence 4, Appl
19	101	36.9	49	2 US-08-694-865-4	Sequence 4, Appl
20	101	36.9	49	2 US-08-878-748-4	Sequence 4, Appl
21	101	36.9	49	3 US-09-124-491-4	Sequence 4, Appl
22	100	36.5	19	1 US-07-610-525-2	Sequence 2, Appl
23	97	35.4	40	4 US-09-026-276-35	Sequence 35, Appl
24	97	35.4	41	4 US-09-026-276-34	Sequence 34, Appl
25	96	35.0	44	1 US-07-690-983D-45	Sequence 45, Appl
26	94.5	34.5	20	2 US-08-319-704-11	Sequence 11, Appl
27	88	32.1	17	1 US-07-610-525-3	Sequence 3, Appl

28	87.5	31.9	22	2 US-08-817-933A-8	Sequence 8, Appl
29	85	31.0	20	1 US-07-690-983D-40	Sequence 40, Appl
30	85	31.0	20	4 US-09-026-276-26	Sequence 26, Appl
31	85	31.0	20	4 US-09-026-276-29	Sequence 29, Appl
32	85	31.0	20	4 US-09-026-276-30	Sequence 30, Appl
33	85	31.0	20	4 US-09-026-276-31	Sequence 31, Appl
34	85	31.0	24	1 US-07-690-983D-43	Sequence 43, Appl
35	79	28.8	14	1 US-07-610-525-6	Sequence 6, Appl
36	79	28.8	14	1 US-08-787-547-43	Sequence 43, Appl
37	79	28.8	15	1 US-07-610-525-4	Sequence 4, Appl
38	79	28.8	15	1 US-08-787-547-44	Sequence 44, Appl
39	71	25.9	17	1 US-08-188-223-6	Sequence 6, Appl
40	71	25.9	17	4 US-08-968-466-6	Sequence 6, Appl
41	71	25.9	17	4 US-08-478-546B-6	Sequence 6, Appl
42	69	25.2	13	1 US-07-610-525-5	Sequence 5, Appl
43	62	22.6	14	1 US-08-453-588-23	Sequence 23, Appl
44	62	22.6	14	3 US-08-521-079-23	Sequence 23, Appl
45	62	22.6	17	1 US-07-690-983D-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-446-692-14
; Sequence 14, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chong Yi
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-14

Query Match 55.5%; Score 152; DB 1; Length 32;
Best Local Similarity 87.9%; Pred. No. 4.4e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 17 FNNFTVSFWLRVPKVSASHLEGLSLHWSYGLRP 49
|||||
Db 3 FNNFTVSFWLRVPKVSASHLE----HWSYGLRP 31

RESULT 2
US-08-488-351A-14
; Sequence 14, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-14
Query Match 55.5%; Score 152; DB 2; Length 32;
Best Local Similarity 87.9%; Pred. No. 4.4e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 1;
QY 17 FNNFTVFWLRVPKVSASHLEPSLHWSYGLRP 49
Db 3 FNNFTVFWLRVPKVSASHLE----HWSYGLRP 31
RESULT 3
PCT-US93-11703-64
; Sequence 64, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green

STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.308
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11703
FILING DATE: 28-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,852
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0222.101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-11703-64
Query Match 41.6%; Score 114; DB 5; Length 31;
Best Local Similarity 95.5%; Pred. No. 6e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 16 LENFTVFWLRVPKVSASHLE 37
Db 6 MENFTVFWLRVPKVSASHLE 27
RESULT 4
US-07-610-525-1
; Sequence 1, Application US/07610525
; Patent No. 5196512
; GENERAL INFORMATION:
; APPLICANT: BIANCHI Eisabetta
; APPLICANT: PESSI Antonello
; APPLICANT: CORRADIN Giampietro
; TITLE OF INVENTION: SYNTHETIC PEPTIDES USEFUL AS
; TITLE OF INVENTION: UNIVERSAL CARRIERS FOR THE PREPARATION OF IMMUNOGENIC COMPOUNDS
; TITLE OF INVENTION: AND THEIR USE IN THE DEVELOPMENT OF SYNTHETIC VACCINES.
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHEA & GOULD
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10020-1193
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/610,525
; FILING DATE: 19901108
; CLASSIFICATION: 424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-827-3000
; TELEFAX: 212-840-6702

TELEX: 423973
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acid residues
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: no
FRAGMENT TYPE: internal fragment
US-07-610-525-1

Query Match 40.9%; Score 112; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
|||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 5
US-08-661-052-12
; Sequence 12, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezman Somsundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISD
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-661-052-12

Query Match 40.9%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
|||||

Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 6
US-08-460-502-8
; Sequence 8, Application US/08460502
; Patent No. 5843464
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gollick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-502-8

Query Match 40.9%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
|||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 7
US-08-724-774B-5
; Sequence 5, Application US/08724774B
; Patent No. 5908778
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
; APPLICANT: Pierref; Cerrottini, Jean-Charles; Carrel,
; APPLICANT: Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
; TITLE OF INVENTION: Rejection Antigen Precursors Mage-10,
; TITLE OF INVENTION: Antibodies Specific To The Molecule, and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,774B
; FILING DATE: 03-October-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5908778man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
US-08-724-774B-5

Query Match 40.9%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 8
US-09-089-595-5
; Sequence 5, Application US/09089595
; Patent No. 6153728
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie, Pierre;
; APPLICANT: Pierre; Carrozzini, Jean-Charles; Carrel, Stefan; Daryl
; APPLICANT: Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
; TITLE OF INVENTION: Rejection Antigen Precursors MAGE-10,
; TITLE OF INVENTION: Antibodies Specific To The Molecule, and
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6153728man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
US-09-089-595-5

Query Match 40.9%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 9
US-09-382-855-5
; Sequence 5, Application US/09382855
; Patent No. 6174692
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie, Pierre;
; APPLICANT: Carrozzini, Jean-Charles; Carrel, Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor Rejection
; TITLE OF INVENTION: Antigen Precursors MAGE-10, Antibodies Specific To The Molec
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/382,855
; FILING DATE: 25-August-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/089,595
; FILING DATE: 02-June-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Schofield, Mary Anne
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
US-09-382-855-5

Query Match 40.9%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10
US-09-183-714B-5
; Sequence 5, Application US/09183714B
; Patent No. 6221593

GENERAL INFORMATION:
APPLICANT: Boon-Falleur, Thierry
APPLICANT: Brasseur, Francis
APPLICANT: Rimoldi, Donata
APPLICANT: De Plaen, Etienne
TITLE OF INVENTION: Method for Determining Cancer by Determining Expression
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/183.714B
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: US 08/724,774
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 7
SEQ ID NO 5
LENGTH: 21
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
US-09-183-714B-5

Query Match 40.9%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSEWLRVPKVSASHLE 37
Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 11
US-09-188-082-12
Sequence 12, Application US/09188082
Patent No. 6270765
GENERAL INFORMATION:
APPLICANT: Yashwant M. Deo
APPLICANT: Joel Goldstein
APPLICANT: Robert Graziano
APPLICANT: Chezian Somasundaram
TITLE OF INVENTION: OF ANTI-PC RECEPTOR ANTIBODIES
TITLE OF INVENTION: OF ANTI-PC RECEPTOR ANTIBODIES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/188.082
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/661,052
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MXI-043CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal

US-09-188-082-12

Query Match 40.9%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSEWLRVPKVSASHLE 37
Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 12
US-09-171-969-10
Sequence 10, Application US/09171969
Patent No. 6284533
GENERAL INFORMATION:
APPLICANT: Thomas, Lawrence J.
TITLE OF INVENTION: PLASMIID-BASED VACCINE FOR TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 75 State Street, Suite 2300
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1807
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171.969
FILING DATE: 01 May 1997 (01.05.97)
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/640,713
FILING DATE: 01 May 1996 (01.05.96)
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/802,967
FILING DATE: 21 February 1997 (21.02.97)
ATTORNEY/AGENT INFORMATION:
NAME: Leon R. Yankwich
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FEATURE:
NAME/KEY:
LOCATION:
US-09-171-969-10

Query Match 40.9%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSEWLRVPKVSASHLE 37
Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 13
PCT-US93-11703-66
Sequence 66, Application PC/TUS9311703
GENERAL INFORMATION:
APPLICANT: Chiron Mimotopes Pty. Ltd.
TITLE OF INVENTION: T-Cell Epitopes

; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-11703-66

Query Match 40.9%; Score 112; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 14
US-08-446-692-5
; Sequence 5, Application US/08446692
; Patent No. 5795551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Y1
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-5

Query Match 39.1%; Score 107; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHL 36
Db 3 FNNFTVSFWLRVPKVSASHL 22

RESULT 15
US-08-488-351A-5
; Sequence 5, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Y1
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-5

Query Match 39.1%; Score 107; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 FNNFTVSFWLRVPKVSASHL 36
Db 3 FNNFTVSFWLRVPKVSASHL 22

Search completed: October 10, 2002, 23:00:55
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 23:02:00 ; Search time 16 Seconds
(without alignments)
276.257 Million cell updates/sec

Title: US-09-848-834A-19
Perfect score: 244
Sequence: 1 XHWSYGLRPGSGPSLQYIK.....GITELSSGPSLHWSYGLRPX 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 10702

Minimum DB seq length: 0
Maximum DB seq length: 46

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	58	23.8	10	1	RHPGG
2	58	23.8	10	1	RHSHG
3	54	22.1	10	1	RHAQ1
4	45	18.4	10	2	A21114
5	40	16.4	10	1	RHAQ2
6	40	16.4	10	1	A61126
7	40	16.4	10	2	A46030
8	40	16.4	10	2	B46030
9	39.5	16.2	45	2	S71834
10	39.5	16.2	45	2	S71835
11	37	15.2	10	2	A49187
12	37	15.2	33	2	PC2300
13	37	15.2	46	2	S77929
14	36.5	15.0	34	2	PH1746
15	36.5	15.0	36	2	PH1751
16	36.5	15.0	44	2	S28770
17	36	14.8	26	2	S00579
18	36	14.8	31	2	PN0050
19	36	14.8	33	2	B82070
20	35.5	14.5	34	2	PH1747
21	35	14.3	44	2	S56313
22	34.5	14.1	33	2	PH1736
23	34.5	14.1	34	2	PH1749
24	34.5	14.1	35	2	S21224
25	34.5	14.1	36	2	PH1752
26	34.5	14.1	36	2	PH1753
27	34.5	14.1	38	2	A55224
28	34	13.9	39	2	H86068
29	34	13.9	40	2	S33407
					gonadoliberin - pi
					gonadoliberin - sh
					gonadoliberin I -
					gonadoliberin II -
					gonadoliberin I - sp
					gonadoliberin I -
					gonadoliberin II -
					cuticular protein
					cuticular protein
					gonatotropin-relea
					gaegurin I - Korea
					exoskeletal protei
					Ig heavy chain v r
					Ig heavy chain v r
					hypothetical prote
					hypothetical prote
					unidentified QM02
					hypothetical prote
					Ig heavy chain v r
					GUR8-2a protein -
					Ig heavy chain v r
					Ig heavy chain v r
					Ig heavy chain v r
					sperm protein EM5
					Ig heavy chain v r
					Ig heavy chain v r
					hypothetical prote
					partial repeat of
					Ig heavy chain v r

30	33.5	13.7	32	2	S20771	Ig heavy chain v r
31	33	13.5	16	2	S57517	T cell receptor be
32	33	13.5	30	2	I49412	gamma-crystallin-3
33	33	13.5	39	2	I65323	calpain II - rat (
34	33	13.5	42	2	E81911	hypothetical prote
35	32.5	13.3	32	2	AH0535	hypothetical prote
36	32.5	13.3	46	4	A45758	hypothetical gluco
37	32	13.1	17	2	I51203	myosin heavy chain
38	32	13.1	35	2	S51708	thioglycosidase (E
39	32	13.1	39	2	A46662	collagen alpha 2(V
40	32	13.1	44	2	S00179	acidic ribosomal p
41	31.5	12.9	33	2	PH1742	Ig heavy chain v r
42	31	12.7	21	2	C39543	collagen alpha 3(I
43	31	12.7	35	2	S26887	Ig heavy chain v r
44	31	12.7	35	2	S46473	Ig heavy chain v r
45	31	12.7	35	2	T42121	hypothetical prote

ALIGNMENTS

RESULT 1
RHPGG
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A>Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A>Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A>Note: the synthetic and natural
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A>Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A>Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
| | | | | | | |
DB 2 HWSYGLRPG 10

RESULT 2
RHSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A>Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780
A:Molecule type: protein

A:Residues: 1-10 <BUR>
A>Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 3
RHAQI
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regal, Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 22.1%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.64;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 4
A21114
gonadoliberin - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match 18.4%; Score 45; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 5
RHAQ2
gonadoliberin II - American alligator
N:Alternate names: gonadotropin-releasing hormone II

C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: B60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regal, Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: B60066
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 16.4%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 6
A61126
gonadoliberin - spotted ratfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Hydrolagus colliei (spotted ratfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C:Accession: A61126
R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.
Gen. Comp. Endocrinol. 82, 152-161, 1991
A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a holostean fish
A:Reference number: A61126; MUID:91340067
A:Accession: A61126
A:Molecule type: protein
A:Residues: 1-10 <LOW>
A:Experimental source: brain
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 16.4%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 7
A46030
gonadoliberin I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Pet-L
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 16.4%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 38;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:| ||

Db 2 HWSHGWLP 10

RESULT 8
B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A>Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 16.4%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:| ||

Db 2 HWSHGWYP 10

RESULT 9
S71834
cuticular protein isoform HACP4.6a - American lobster
C:Species: Homarus americanus (American lobster)
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 25-Apr-1997
C:Accession: S71834
R:Kragh, M.; Molbak, L.; Andersen, S.O.
submitted to the Protein Sequence Database, November 1996
A:Description: Cuticular proteins from the lobster, Homarus americanus.
A:Reference number: S71832
A:Accession: S71834
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-45 <KRA>
F:1-45/Product: cuticular protein isoform HACP4.6a #status experimental <MAT>

Query Match 16.2%; Score 39.5; DB 2; Length 45;
Best Local Similarity 41.7%; Pred. No. 2.5e+02;
Matches 10; Conservative 3; Mismatches 6; Indels 5; Gaps 1;

QY 13 GPSLOVIKANSKFITELSSGPS 36
|:|: || | ||| |||

Db 13 GTQVQFSKA-----GVTVLLEGPS 31

RESULT 10
S71835
cuticular protein isoform HACP4.6b - American lobster
C:Species: Homarus americanus (American lobster)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C:Accession: S71835
R:Kragh, M.; Molbak, L.; Andersen, S.O.
submitted to the Protein Sequence Database, November 1996
A:Description: Cuticular proteins from the lobster, Homarus americanus.
A:Reference number: S71832
A:Accession: S71835
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-45 <KRA>

F:1-45/Product: cuticular protein isoform HACP4.6b #status experimental <MAT>

Query Match 16.2%; Score 39.5; DB 2; Length 45;
Best Local Similarity 41.7%; Pred. No. 2.5e+02;
Matches 10; Conservative 3; Mismatches 6; Indels 5; Gaps 1;

QY 13 GPSLOVIKANSKFITELSSGPS 36
|:|: || | ||| |||

Db 13 GTQVQFSKA-----GVTVLLEGPS 31

RESULT 11
A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A>Title: Primary structure and biological activity of a third gonadotropin-releasing
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A:Molecule type: protein
A>Status: preliminary
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A>Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 15.2%; Score 37; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||: ||

Db 2 HWSHDKWPG 10

RESULT 12
PC2300
gaegurin 1 - Korean frog (fragment)
C:Species: Rana rugosa (Korean frog)
C>Date: 25-Feb-1995 #sequence_revision 26-May-1995 #text_change 19-May-2000
C:Accession: PC2300
R:Park, J.M.; Jung, J.E.; Lee, B.J.
Biochem. Biophys. Res. Commun. 205, 948-954, 1994
A>Title: Antimicrobial peptides from the skin of a Korean frog, Rana rugosa.
A:Reference number: PC2300; MUID:95091844
A:Accession: PC2300
A:Molecule type: protein
A:Residues: 1-33 <PAR>
C:Superfamily: ranalexin precursor; dermorphin precursor amino-terminal homology
C:Keywords: skin
F:27-33/Region: rana box motif

Query Match 15.2%; Score 37; DB 2; Length 33;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 19 IKANSKFITELSSG 34
|||: |||: |

Db 6 IKAGAKFLCKNLLKQG 21

RESULT 13
S77929
exoskeletal protein HACP44 - American lobster
C:Species: Homarus americanus (American lobster)
C>Date: 23-Jul-1997 #sequence_revision 01-Aug-1997 #text_change 01-Aug-1997
C:Accession: S77929
R:Nousiainen, M.; Rafn, K.; Skou, L.; Roepstorff, P.; Andersen, S.O.
submitted to the Protein Sequence Database, June 1997
A:Description: Characterization of exoskeletal proteins from the American lobster, Ho
A:Reference number: S77925

Fri Oct 11 06:16:19 2002

A;Accession: S77929
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-46 <NOU>

Query Match	15.2%;	Score 37;	DB 2;	Length 46;
Best.Local Similarity	40.0%;	Pred. No. 5.2e+02;		
Matches	12;	Conservative	6;	Mismatches 6;
		Indels	6;	Gaps 2;

QY	10	GSSG--	--PSLQYIKANSKFIGITELSSG	36
		: :	: : : : : : : :	
Db	7	GASGIITPSGRLLQLPA--	--GVTVASAGPS	33

RESULT 14

PH1746
Ig heavy chain V region (clone NP-12-12) - mouse (fragment)
C.Species: Mus musculus (house mouse)
C.Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C.Accession: PH1746
R.R.McMeyer-Williams, M.G.: McLean, M.J.: Lalor, P.A.: Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A.Title: Antigen-driven B cell differentiation in vivo.
A.Reference number: PH1675; MUID:93301607

Query Match	15.0%;	Score 36.5;	DB 2;	Length 34;
Best Local Similarity	32.5%;	Pred. No. 4.3e+02;		
Matches 13;	Conservative 3;	Mismatches 13;	Indels 11;	Gaps 2;

3 WSYGLRPGS-SGPSLQYIKANSKFITELSSGSPSLHWSY 41
5 WVKGREPGTGTSKPSMO-----STERYTGSSSHFDY 34

RESULT 15

PH1751
Ig heavy chain V region (clone NP-12-17) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1751
R:McMeyer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607
A:Accession: PH1751
A:Molecule type: mRNA
A:Residues: 1-36 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 15.0%; Score 36.5; DB 2; Length 36;
Best Local Similarity 30.0%; Pred. No. 4.6e+02;
Matches 12; Conservative 4; Mismatches 15; Indels 9; Gaps 2;

```

3 WSYGLRPGS-SGPSLQYKANSKFIGITELSGPSSLHSY 41
  | | | | | | | | | | | | | | | | | | |
5 WVKRPGTGKSPKPSMSTERYHY-----GSSCHGDY 36
  | | | | | | | | | | | | | | | | | | |

```

Search completed: October 10, 2002, 23:06:28
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 23:00:25 ; Search time 13 Seconds
(without alignments)
137.008 Million cell updates/sec

Title: US-09-848-834A-19

Perfect score: 244

Sequence: 1 XHWSYGLRPGSSGFSLOYIK.....GITELSSGSLHWSYGLRPX 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 3308

Minimum DB seq length: 0

Maximum DB seq length: 46

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	54	22.1	10	1 GONI_ALLMI	P37041 alligator m
2	47	19.3	10	1 GONI_CLUPA	P81749 Clupea pall
3	45	18.4	10	1 GONI_ONCKE	P20367 oncorhynch
4	40	16.4	10	1 GONI_CHICK	P37043 gallus gall
5	40	16.4	10	1 GONI_SQUAC	P27429 squallus aca
6	38.5	15.8	46	1 GONI_PETMA	P82889 bacterioph
7	37	15.2	10	1 COAT_BPH75	P30948 petromyzon
8	37	15.2	33	1 GAE1_RANRU	P80395 rana rugosa
9	37	15.2	33	1 RUBG_RANRU	P80955 rana rugosa
10	34.5	14.1	36	1 HIL5_ENSMI	P27203 ensis minor
11	33.5	13.7	39	1 ABAB_BOMPA	P81463 bombus pasc
12	33	13.5	43	1 NT3_RAJCL	P25434 raja clav
13	33	13.5	44	1 PIS1_CARPI	P80569 carnobacter
14	32	13.1	10	1 GONI_CHEPR	P80677 chelyosoma
15	32	13.1	43	1 MUT1_ENTMU	P80925 enterococcu
16	32	13.1	44	1 RLA2_RABIT	P19943 eryctolagus
17	30	12.3	37	1 GRP_CHICK	P01295 gallus gall
18	30	12.3	32	1 MIFH_TRITR	P81748 trichuris t
19	30	12.3	44	1 GG12_STAHA	P11698 staphylococ
20	29.5	12.1	43	1 KAD_CORGL	P49973 corynebacte
21	29	11.9	28	1 GUN_SCHCO	P81190 schizophyll
22	29	11.9	36	1 NPF_ARTTR	P41334 artiposthi
23	29	11.9	38	1 MUTS_YEREN	P47763 yersinia en
24	29	11.9	43	1 PSBN_MESVI	Q9muv5 mesostigma
25	29	11.9	44	1 BGAL_PINPS	P81669 pinus pinas
26	29	11.9	44	1 GG11_STAHA	P11697 staphylococ
27	28.5	11.7	27	1 GRP_CANFA	P08989 canis fami
28	28.5	11.7	29	1 PSBF_SYN6	P31083 synchococc
29	28	11.5	10	1 GONI_CHEPR	P80678 chelyosoma
30	28	11.5	21	1 NRLA_ACTSP	P33036 acinetobact
31	28	11.5	36	1 TAEK_ACTEO	P81897 actinia equ
32	28	11.5	44	1 CPC9_CANPG	P81590 cancer pagu
33	28	11.5	46	1 TXA5_ANESU	P01529 anemonia su

34 27 11.1 18 1 SODM_MYCHA P80582 mycobacteri
35 27 11.1 27 1 PSBF_SYN6 P23078 synchococc
36 27 11.1 33 1 Y849_BORBU O51789 borrelia bu
37 27 11.1 36 1 LCCC_LEUME P81053 leuconostoc
38 27 11.1 37 1 PETG_EUGGR P30396 euglena gra
39 27 11.1 43 1 PSBN_PORPU P51324 porphyra pu
40 27 11.1 46 1 DIUH_LOCHI P23465 locusta mig
41 27 11.1 46 1 DIUH_PERAM P41538 periplaneta
42 26.5 10.9 19 1 TRP3_LEUMA P81735 leucophaea
43 26.5 10.9 27 1 TXA3_ANESU P01535 anemonia su
44 26.5 10.9 40 1 YPS1_PLEBO P15733 plectonema
45 26.5 10.9 43 1 BDNF_RAJCL P25430 raja clav

ALIGNMENTS

RESULT 1
GONI_ALLMI STANDARD; PRT; 10 AA.
ID GONI_ALLMI P37041; P20407;
AC 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC -|- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -|- SURCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;
Query Match 22.1%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.22;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 2 HWSYGLQPG 10
RESULT 2
GONI_CLUPA STANDARD; PRT; 10 AA.
ID GONI_CLUPA P81749;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH)
DE (Luliberin I).
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;

Qy 2 HWSYGLRPG 10
| | | : | |
Db 2 HWSHDWPG 10

RESULT 8

ID GAEL_RANRU STANDARD; PRT; 33 AA.
AC P80395;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gaegurin-1.
OS Rana rugosa (Wrinkled frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=8410;
RN [1]

TISSUE=Skin;

RX MEDLINE=95091844; PubMed=7999137;

RA Park J.M., Jung J.-P., Lee B.J.;

RT "Antimicrobial peptides from the skin of a Korean frog, Rana

rugosa.";

RL Biochem. Biophys. Res. Commun. 205:948-954(1994).

CC -1- FUNCTION: HAS A NON-HEMOLYTIC ACTIVITY. HAS A BROAD SPECTRUM OF

CC ACTIVITY AGAINST BOTH GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA,

CC FUNGI AND PROTOZOA.

CC -1- SUBUNIT: MONOMER.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: SKIN.

CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN

CC FAMILY.

KW Amphibian skin; Antibiotic.

FT DI6ULFID 27 33 BY SIMILARITY.

SQ SEQUENCE 33 AA; 3462 MW; 51E21E3B7B0FF536 CRC64;

Query Match 15.2%; Score 37; DB 1; Length 33;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 19 IKANSKFIGITELSSG 34
| | | : | | | |
Db 6 IKAGAKFLGKNLLKQG 21

RESULT 9

ID RUGB_RANRU STANDARD; PRT; 33 AA.
AC P80955;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Rugosin B.
OS Rana rugosa (Wrinkled frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=8410;
RN [1]

TISSUE=Skin;

RX MEDLINE=95336450; PubMed=7612013;

RA Suzuki S., Ohe Y., Kagegawa T., Tatemoto K.;

RT "Isolation and characterization of novel antimicrobial peptides,

RT rugosins A, B and C, from the skin of the frog, Rana rugosa.";

RL Biochem. Biophys. Res. Commun. 212:249-254(1995)

CC -1- FUNCTION: SHOWS ANTIBACTERIAL ACTIVITY AGAINST BOTH GRAM-NEGATIVE

CC AND GRAM-POSITIVE BACTERIA.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: SKIN.

CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN

CC FAMILY.

KW Antibiotic; Amphibian skin.

FT DISULFID 27 33
SQ SEQUENCE 33 AA; 3516 MW; 51FFC8E1641FF536 CRC64;

Query Match 15.2%; Score 37; DB 1; Length 33;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 19 IKANSKFIGITELSSG 34
| | | : | | | |
Db 6 IKAGAKFLGKNLLKQG 21

RESULT 10

ID HIL5_ENSMI STANDARD; PRT; 36 AA.
AC P27203;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Histone H1-like protein EM5 (Fragment).
OS Ensis minor (Razor shell) (Minor jackknife clam).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Solenoidea; Pharidae; Ensis.
OX NCBI_TaxID=6587;
RN [1]

TISSUE=Sperm;

RX MEDLINE=92130249; PubMed=1547275;

RA Giannocotti V., Buratti E., Santucci A., Neri P., Crane-Robinson C.;

RT "Molluscan sperm proteins: Ensis minor.";

RL Biochim. Biophys. Acta 1119:296-302(1992).

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- TISSUE SPECIFICITY: SPERM.

CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.

DR PIR; S21224; S21224.

DR HSSP; P02259; IHST.

DR InterPro; IPR001386; Linker_histone.

DR Pfam; PF00538; linker_histone; 1.

DR SMART; SM00526; H15; 1.

KW Chromosomal protein; Nuclear protein; DNA-binding.

FT NON_TER 1 1 GLOBULAR.

FT DOMAIN 1 36

FT NON_TER 36 36

SQ SEQUENCE 36 AA; 3774 MW; A5EA2ECC6C88412 CRC64;

Query Match 14.1%; Score 34.5; DB 1; Length 36;
Best Local Similarity 55.6%; Pred. No. 3.1e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Qy 8 RPSGGPS-LQYIKANSK 24
| | | | : | | | | |
Db 12 RGGSSRQAILKYIQANFK 29

RESULT 11

ID ABAAE_BOMPA STANDARD; PRT; 39 AA.
AC P81463;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Abaecin.
OS Bombus pascuorum.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Aculeata;
OC Apoidea; Apidae; Bombus.
OX NCBI_TaxID=65598;
RN [1]

SEQUENCE.

RX MEDLINE=9219367;

RA Rees J.A., Monlatte M., Bulet P.;

RT "Novel antibacterial peptides isolated from a European bumblebee,

Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Carnobacterium.
OX NCBI_TaxID=2751.
RN [1]
RP SEQUENCE.
RC STRAIN=JG126;
RX MEDLINE=96316406; PubMed=8702282;
RJ Jack R.W., Wan J., Gordon J., Harmark K., Davidson B.E., Hillier A.J.,
RK Mettenhall R.E.H., Hickey M.W., Coventry J.M.;
RL "Characterization of the chemical and antimicrobial properties of
RT piscicolin 126, a bacteriocin produced by Carnobacterium piscicola
RJG126".
RT Appl. Environ. Microbiol. 62:2897-2903(1996).
RL -1- FUNCTION: INHIBITS THE GROWTH OF SEVERAL GRAM-POSITIVE BACTERIA,
CC ESPECIALLY THE FOOD-BORNE PATHOGEN LISTERIA MONOCYTOGENES, BUT HAS
CC NO EFFECT ON THE GROWTH OF A NUMBER OF YEASTS AND GRAM-NEGATIVE
CC BACTERIA.
CC -1- MASS SPECTROMETRY: MW=4416.6; MW_ERR=1.9; METHOD-Electrospray.
CC -1- SIMILARITY: BELONGS TO THE "YGNCV" FAMILY OF BACTERIOCINS.
CC -1- MASS SPECTROMETRY: MW=4418 MW; 858608B15EE72568 CRC64;
DR HSP; P34034; 2LEU.
DR InterPro; IPR002633; Bacteriocin_II.
DR Pfam; PF01721; Bacteriocin_II; 1.
DR ProDom; PD004452; Bacteriocin_II; 1.
KW Antibiotic; Bacteriocin.
RT DISULFID 9 14 BY SIMILARITY.
FT SEQUENCE 44 AA; 4418 MW; 858608B15EE72568 CRC64;
SQ

Query Match 13.5%; Score 33; DB 1; Length 44;
Best Local Similarity 23.4%; Pred. No. 6e+02;
Matches 11; Conservative 11; Mismatches 15; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGSLQYKANSKFIGI-----TELSSGPSLHWSYG 42
: : | : | : : | | | | : : | : | : |
DB 2 YYGNGVSCKNGCTVDW----SKAIIIGNNAANLTGTGAAGWNKG 44

RESULT 14
GONI_CHEPR STANDARD; PRT; 10 AA.
ID ID GONI_CHEPR
AC P80677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I)
DE (Luiberlin I).
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Corellidae; Chelyosoma.
OC NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
RT and the evolutionary implications";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUPEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
CC THE GONADUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -1- MASS SPECTROMETRY: MW=1246.56; METHOD-MALDI.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro; IPR002012; GNRH.
KW PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
SQ SEQUENCE 10 AA; 1264 MW; 284B3639DB5AB5A3 CRC64;
Query Match 13.1%; Score 32; DB 1; Length 10;

Search completed: October 10, 2002, 23:05:02
Job time : 14 secs

;
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 23:01:00 ; Search time 28 Seconds
(without alignments)
284.206 Million cell updates/sec

Title: US-09-848-834A-19

Perfect score: 244

Sequence: 1 XHNSYGLRPGSSGPSLOYIK.....GITELSSGSLHWSYGLRPX 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 26879

Minimum DB seq length: 0

Maximum DB seq length: 46

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	18.4	33	13 Q9PT34	Q9pt34 oncorhynch
2	45	18.4	33	13 Q9W7G0	Q9w7g0 oncorhynch
3	44	18.0	40	12 Q91K37	Q91k37 hepatitis c
4	42	17.2	36	8 Q951P2	Q951p2 salmo salar
5	42	17.2	36	8 Q94VL4	Q94vl4 salmo trutt
6	42	17.2	36	8 Q94NV5	Q94ny5 salmo salar
7	42	17.2	39	8 Q34493	Q34493 equus quagg
8	42	17.2	39	8 Q33434	Q33434 equus burch
9	41	16.8	35	8 Q951P1	Q951p1 salmo salar
10	40	16.4	43	3 Q92217	Q92217 saccharomyc
11	39	16.0	40	12 Q91K61	Q91k61 hepatitis c
12	39	16.0	40	12 Q91K60	Q91k60 hepatitis c
13	39	16.0	42	3 Q42780	Q42780 ustilago ho
14	38	15.6	20	11 Q9EOX8	Q9eox8 mus musculu
15	38	15.6	32	5 Q9U340	Q9u340 caenorhabdi
16	37	15.2	35	15 Q9QQI5	Q9qqiy5 human Immun

17	37	15.2	40	4 Q9BW55	Q9bw55 homo sapien
18	37	15.2	45	4 Q96IQ2	Q96iq2 homo sapien
19	36.5	15.0	44	8 Q32439	Q32439 hordeum vul
20	36	14.8	26	1 Q50835	Q50835 methanococc
21	36	14.8	30	4 Q9UMI6	Q9umi6 homo sapien
22	36	14.8	33	16 Q9KP87	Q9kp87 vibrio chol
23	36	14.8	45	8 Q78734	Q78734 flammulina
24	35.5	14.5	45	2 Q9EX34	Q9ex34 streptomyce
25	34.5	14.1	27	12 Q91J64	Q91j64 hepatitis c
26	34.5	14.1	27	12 Q91J63	Q91j63 hepatitis c
27	34.5	14.1	27	12 Q91J62	Q91j62 hepatitis c
28	34.5	14.1	27	12 Q91J61	Q91j61 hepatitis c
29	34.5	14.1	38	2 Q53057	Q53057 rhodobacter
30	34	13.9	33	13 Q9W7N1	Q9w7n1 morone saxa
31	34	13.9	35	2 Q9R4S7	Q9r4s7 rhodobacter
32	34	13.9	39	8 Q33435	Q33435 equus cabal
33	34	13.9	45	2 Q9LBU2	Q9lbu2 escherichia
34	33	13.5	27	12 Q68906	Q68906 hepatitis c
35	33	13.5	30	11 Q62536	Q62536 mus spretus
36	33	13.5	31	2 Q9R9C1	Q9r9cl borrella bu
37	33	13.5	33	12 Q9PXD3	Q9pxd3 hepatitis c
38	33	13.5	36	8 Q951P0	Q951p0 salmo salar
39	33	13.5	42	16 Q9JUA3	Q9jua3 neisseria m
40	32.5	13.3	20	11 Q9QVF6	Q9qvfe rattus sp.
41	32.5	13.3	41	5 Q917S0	Q917s0 drosophila
42	32	13.1	17	13 Q91352	Q91352 gallus gall
43	32	13.1	24	10 Q9S937	Q9s937 beta vulgar
44	32	13.1	27	12 Q39932	Q39932 hepatitis c
45	32	13.1	31	5 Q9TWK5	Q9twk5 mytilus edu

ALIGNMENTS

RESULT 1

Q9PT34	ID	Q9PT34	PRELIMINARY;	PRT;	33 AA.
AC	Q9PT34;				
DC	01-MAY-2000 (TREMBlrel. 13, Created)				
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)				
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)				
DE	GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)				
DE	(LULIBERIN) (FRAGMENT).				
GN	GNRH.				
OS	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;				
OC	Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.				
OX	NCBI_TaxID=8022;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99312119; PubMed=10385393;				
RA	Von Schalburg K.R., Sherwood N.M.;				
RT	"Regulation and expression of gonadotropin-releasing hormone gene				
RT	differs in brain and gonads in rainbow trout.";				
RL	Endocrinology 140:3012-3024(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Von Schalburg K.R., Sherwood N.M.;				
RL	Submitted (DSC-1998) to the EMBL/GenBank/DBJ databases.				
CC	-!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY				
CC	SIMILARITY).				
CC	-!- SIMILARITY: BELONGS TO THE GNRH FAMILY.				
DR	EMBL; AF110533; AAD43461.1; -.				
DR	InterPro; IPR002047; AKH.				
DR	InterPro; IPR002012; GNRH.				
DR	Pfam; PF00446; GNRH; 1.				
DR	PROSITE; PS00256; AKH; UNKNOWN_1.				
DR	PROSITE; PS00473; GNRH; 1.				
KW	Amidation; Hormone.				
FT	NON_TER 33				
FT	SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;				


```
OS Salmo salar (Atlantic salmon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V62;
RA Consuegra S., Garcia de Leaniz C., Sordio A.L., Knox D., Straus L.G.,
RA Gonzalez-Morales M.R., Verspoor E.;
RT "Mitochondrial DNA variation in Paleolithic and modern Atlantic salmon
from the Iberian glacial refugium."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF385756; AAK67351.1; -
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 35
SQ SEQUENCE 35 AA; 3744 MW; B67719832E2E597F CRC64;
Query Match 16.8%; Score 41; DB 8; Length 35;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 24 KFIGITELSSGSPSLHWSYGL 43
I : I : I : I : I :
D 8 KVLGXXLXKGNIVGPGYGL 27
RESULT 10
Q02217 PRELIMINARY; PRT; 43 AA.
AC Q02217;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TIF51A PROTEIN (FRAGMENT).
GN HYP2 OR TIF51A OR YEL034W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94016558; PubMed=8411151;
RA Melnick L., Sherman F.;
RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
of Saccharomyces cerevisiae share a common ancestry."
RL J. Mol. Biol. 233:372-388(1993).
DR EMBL; S65964; AADI3969.1; -
DR SGD; S0000760; HYP2.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 43 AA; 4728 MW; 65D24A32935EC7B3 CRC64;
Query Match 16.4%; Score 40; DB 3; Length 43;
Best Local Similarity 57.1%; Pred. No. 3.2e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 29 TELSSGSPSLHWSYGL 42
I : I : I : I : I :
D 14 TVLSQGRTLHWVG 27
RESULT 11
Q01K61 PRELIMINARY; PRT; 40 AA.
AC Q01K61;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
```

```
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LC13;
RA Gimenez-Barcons M., Franco S., Suarez Y., Forns X., Ampurdanes S.,
RA Puig-Basagoiti F., Sanchez-Fueyo A., Barrera J.M., Llovet J.M.,
RA Bruix J., Sanchez-Tapias J.M., Rodes J., Salz J.C.;
RT "High amino acid variability within the NS5A gene of hepatitis C virus
(HCV) is associated with hepatocellular carcinoma in patients with
HCV-1b related cirrhosis."
RT Hepatology 0:0-0(2001).
DR EMBL; AF379211; AAK63428.1; -
FT NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4233 MW; 42366445DBF75C0C CRC64;
Query Match 16.0%; Score 39; DB 12; Length 40;
Best Local Similarity 29.6%; Pred. No. 4e+02;
Matches 8; Conservative 7; Mismatches 8; Indels 4; Gaps 1;
QY 21 ANSKFIGI----TELSSGSPSLHWSYGL 43
I : I : I : I : I : I :
D 1 ARASFGVDGXATRVGTGTAHTTYRL 27
RESULT 12
Q01K60 PRELIMINARY; PRT; 40 AA.
AC Q01K60;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LC14;
RA Gimenez-Barcons M., Franco S., Suarez Y., Forns X., Ampurdanes S.,
RA Puig-Basagoiti F., Sanchez-Fueyo A., Barrera J.M., Llovet J.M.,
RA Bruix J., Sanchez-Tapias J.M., Rodes J., Salz J.C.;
RT "High amino acid variability within the NS5A gene of hepatitis C virus
(HCV) is associated with hepatocellular carcinoma in patients with
HCV-1b related cirrhosis."
RT Hepatology 0:0-0(2001).
DR EMBL; AF379212; AAK63429.1; -
FT NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4233 MW; 42366445DBF75C0C CRC64;
Query Match 16.0%; Score 39; DB 12; Length 40;
Best Local Similarity 29.6%; Pred. No. 4e+02;
Matches 8; Conservative 7; Mismatches 8; Indels 4; Gaps 1;
QY 21 ANSKFIGI----TELSSGSPSLHWSYGL 43
I : I : I : I : I : I :
D 1 ARASFGVDGXATRVGTGTAHTTYRL 27
RESULT 13
Q02780 PRELIMINARY; PRT; 42 AA.
AC Q02780;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MAT-1 PHEROMONE.
GN MFA1.
OS Ustilago hordei (Smut fungus).
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
```

OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
RX NCBI_TaxID=120017;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=14A;
RX MEDLINE=20033551; PubMed=10564815;
RA Anderson C.M., Willis D.A., Koster P.J., Ford E.J.,
RA Martinez-Espinosa A.D., Sherwood J.E.;
RT "Molecular analysis of the pheromone and pheromone receptor genes of
Ustilago hordei.";
RL Gene 240:89-97(1999).
DR EMBL; AF043940; AAC02682.1; -;
SQ SEQUENCE 42 AA; 4215 MW; 7AE171D5E8759BA2 CRC64;

Query Match 16.0%; Score 39; DB 3; Length 42;
Best Local Similarity 46.7%; Pred. No. 4.2e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 SYGLRPGSGSPSQY 18
::: |||||

Db 22 NHGANPKSGSLGY 36

RESULT 14

Q9EQX8 PRELIMINARY; PRT; 20 AA.
AC Q9EQX8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GLUTAMATE DECARBOXYLASE 65 (FRAGMENT).
GN GAD65.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=129SV; TISSUE=LIVER;
RA Makinae K., Kobayashi T., Kobayashi T., Sinkawa H., Sakagami H.,
RA Kondo H., Tashiro F., Miyazaki J., Obata K., Tamura S., Yanagawa Y.;
RT "Structure of the Mouse Glutamate Decarboxylase 65 Gene and its
Promoter.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032757; BAB20415.1; -;
FT NON_TER 20 20

SQ SEQUENCE 20 AA; 1989 MW; 28F544B0E4C3768D CRC64;

Query Match 15.6%; Score 38; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 WSYGLRPGSSGP 14
||:| ||:

Db 9 WSPGSEDGSADP 20

RESULT 15

Q9U340 PRELIMINARY; PRT; 32 AA.
AC Q9U340;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE W06G6.9 PROTEIN.
GN W06G6.9.
OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

RA McMurray A.A.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z83129; CAB63325.1; -;
SQ SEQUENCE 32 AA; 3775 MW; E49D031288C6806A CRC64;

Query Match 15.6%; Score 38; DB 5; Length 32;
Best Local Similarity 41.7%; Pred. No. 4.2e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSG 13
||:| ||:

Db 18 HWQWAAKPGEWG 29

Search completed: October 10, 2002, 23:05:51
Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 22:59:45 ; Search time 27 Seconds
(without alignments)
189.237 Million cell updates/sec

Title: US-09-848-834A-19
Perfect score: 244
Sequence: 1 XHWSYGLRPGSGSLQYIK.....GITELSSGPSLHWSYGLRPX 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 340140

Minimum DB seq length: 0
Maximum DB seq length: 46

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

```

1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	113.5	46.5	27 15 AAR62701	LHRH-containing im
2	94	38.5	42 21 AAB20865	GnRH tandem dimer
3	92.5	37.9	23 21 AAB20864	GnRH tandem repeat
4	91	37.3	30 11 AAR07323	Luteinising hormon
5	91	37.3	40 20 AAY31183	Ubiquitin fusion p
6	91	37.3	40 22 AAB71954	GnRH sequence #2.
7	91	37.3	41 22 AAY31182	Ubiquitin fusion p
8	91	37.3	41 22 AAB71953	GnRH sequence #1.
9	90	36.9	40 21 AAY96085	Cattle gonadotropi
10	87	35.7	20 19 AAW47438	Antigenic peptide.
11	87	35.7	20 20 AAY31174	Ubiquitin fusion p

12	87	35.7	20 20 AAY31177	Ubiquitin fusion p
13	87	35.7	20 20 AAY31178	Ubiquitin fusion p
14	87	35.7	20 20 AAY31179	GnRH dimer. Unide
15	87	35.7	20 22 AAB71945	GnRH dimer peptide
16	87	35.7	20 22 AAB71948	GnRH mixed dimer p
17	87	35.7	20 22 AAB71949	GnRH mixed dimer p
18	87	35.7	20 22 AAB71950	Luteinising hormon
19	87	35.7	21 11 AAR07324	Anyloid beta/tetan
20	80	32.8	27 22 AAB49074	Anyloid beta/tetan
21	80	32.8	27 22 AAB49077	Anyloid beta/tetan
22	76	31.1	19 21 AAY99055	HLA class II bindi
23	76	31.1	37 15 AAR65389	Universal immunost
24	76	31.1	37 15 AAR65383	Universal immunost
25	75.5	30.9	42 18 AAW21649	Peptide containing
26	75	30.7	25 21 AAY92650	PSMpep007 - P2 ins
27	75	30.7	27 16 AAR82596	IgE CH4 region con
28	75	30.7	44 18 AAW21650	Peptide containing
29	74.5	30.5	31 21 AAY82632	Tetanus toxoid T c
30	74	30.3	15 11 AAR06310	Tetanus toxin epit
31	74	30.3	15 18 AAW35506	Universal T-cell e
32	74	30.3	15 18 AAW11505	Tetanus toxoid uni
33	74	30.3	15 19 AAW67033	Tetanus toxin frag
34	74	30.3	15 19 AAW71321	Universal helper T
35	74	30.3	15 20 AAY04051	T-Helper epitope f
36	74	30.3	15 20 AAW67578	T-cell epitope pep
37	74	30.3	15 20 AAW73220	Tetanus toxoid epi
38	74	30.3	15 21 AAB45511	Tetanus P2 epitope
39	74	30.3	15 21 AAY82637	Tetanus toxoid T c
40	74	30.3	15 21 AAY92625	Foreign epitope P2
41	74	30.3	15 21 AAY84427	Amino acid sequenc
42	74	30.3	15 21 AAY70300	Clostridium tetani
43	74	30.3	15 21 AAY44763	Tetanus toxoid pro
44	74	30.3	15 22 AAE11763	Clostridium tetani
45	74	30.3	15 22 AAM99515	Vaccine related MH

ALIGNMENTS

RESULT 1
AAR62701
ID AAR62701 standard; peptide; 27 AA.

XX AAR62701;

AC AAR62701;

XX 10-SEP-1995 (first entry)

DT LHRH-containing immunogenic peptide.

DE Helper T cell epitope; universal immune stimulator; invasin; hapten;

XX KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;

XX KW androgen-dependent carcinoma; antitumour; infertility;

XX KW tetanus toxin.

OS Synthetic.

XX Key Location/Qualifiers

FT Domain 1..17

FT Domain /note= "tetanus toxin helper T cell epitope"

FT Domain 18..27

FT Domain /note= "LHRH haptent"

XX WO9425060-A.

PN 10-NOV-1994.

XX 28-APR-1994; 94WO-US04832.

XX 27-APR-1993; 93US-0057166.

PR 14-APR-1994; 94US-0229275.

PA (LADD/) LADD A E.

PA (WANG/) WANG C Y.

PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX
 XX Claims 8, 12; Page 84; 213pp; English.
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptan containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasive and Th domains and between the immune stimulator and haptan
 CC components. When the haptan is LHRH, then optionally the invasive domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasive-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX This sequence is particularly preferred.

XX Sequence 27 AA;
 Query Match 46.5%; Score 113.5; DB 15; Length 27;
 Best Local Similarity 79.3%; Pred. No. 7.6e-08;
 Matches 23; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
 Qy 17 QYKANSKFIGITELSGPSLHWSYGLRP 45
 Db 3 QYKANSKFIGITELSGPSLHWSYGLRP 26
 |||||
 |||||

RESULT 2
 AAB20865
 ID AAB20865 standard; peptide; 42 AA.

XX AC AAB20865;
 XX DT 03-JAN-2001 (first entry)
 XX GnRH tandem dimer peptide sequence SEQ ID NO:3.
 XX Gonadotropin releasing hormone; GnRH; immunogen; Protein D; carrier;
 KW prostate cancer; Haemophilus influenzae; vaccine; infectious disease;
 KW malaria; cytostatic; antiallergic; neurotropic; neuroprotective;
 KW protozoacide; Alzheimer's disease; allergy.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 42 /note= "amidated"
 FT
 XX

XX WO2000050077-A1.

XX 31-AUG-2000.

XX 22-FEB-2000; 2000WO-EP01457.

XX 25-FEB-1999; 99GB-0004405.

XX 25-FEB-1999; 99GB-0004408.

XX 25-FEB-1999; 99GB-0004412.

PR 13-AUG-1999; 99GB-0019260.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Coste M, Lobet Y, Van-Mechelen MP, Verriest C;
 XX WPI; 2000-572040/53.
 DR Immunogens and vaccine comprising the immunogen useful for preventing
 PT and treating infectious diseases e.g. malaria and chronic disease e.g.
 PT cancer, comprises peptide and carrier from protein D of influenzae
 XX Disclosure; Page 7; 53pp; English.

XX The present invention describes an immunogen (I) comprising a peptide
 CC (Ia) and a carrier (Ib) derived from protein D of Haemophilus influenzae
 CC or its fragment. Also described are: (1) a vaccine comprising (I), and
 CC an excipient; (2) preparation of (I), comprising conjugating a peptide
 CC to protein D or its fragment; and (3) preparation of a vaccine of (1),
 CC comprising formulating (I) with an excipient. (I) has cytostatic,
 CC antiallergic, neurotropic, neuroprotective and protozoacide activities.
 CC (I) and the vaccine are useful for the manufacture of a medicament for
 CC preventing and treating infectious diseases such as malaria or chronic
 CC disease such as cancer, Alzheimer's disease or allergy in a patient.
 CC Unlike prior art immunogens, (I) induces high levels of antipeptide
 CC immune responses while inducing a moderate humoral response against the
 CC carrier. The present sequence represents an example of an immunogen from
 CC the present invention which contains gonadotropin releasing hormone
 CC (GnRH) tandem dimers.

XX Sequence 42 AA;

XX Query Match 38.5%; Score 94; DB 21; Length 42;
 Best Local Similarity 45.7%; Pred. No. 4.3e-05;
 Matches 21; Conservative 0; Mismatches 7; Indels 18; Gaps 2;
 Qy 2 HWSYGLRPGSGPSLQYKANSKFIGITELSGS--PSLHWSYGLRP 45
 Db 12 HWSYGLRPGS-----CBHWSYGLRPGQHWSYGLRP 41
 |||||
 |||||

RESULT 3
 AAB20864

ID AAB20864 standard; peptide; 23 AA.

XX AC AAB20864;

XX DT 03-JAN-2001 (first entry)

XX GnRH tandem repeat peptide sequence SEQ ID NO:2.

XX Gonadotropin releasing hormone; GnRH; immunogen; Protein D; carrier;
 KW prostate cancer; Haemophilus influenzae; vaccine; infectious disease;
 KW malaria; cytostatic; antiallergic; neurotropic; neuroprotective;
 KW protozoacide; Alzheimer's disease; allergy.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 23 /note= "amidated"
 FT
 XX

XX WO2000050077-A1.

XX 31-AUG-2000.

XX 22-FEB-2000; 2000WO-EP01457.

XX 25-FEB-1999; 99GB-0004405.

XX 25-FEB-1999; 99GB-0004408.

XX 25-FEB-1999; 99GB-0004412.

XX 13-AUG-1999; 99GB-0019260.

```

PA (SWIK)_SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX
PI Coste M, Lobet Y, Van-Mechelen MP, Verriest C;
XX
XX WPI; 2000-572040/53.
XX
XX Immunogens and vaccine comprising the immunogen useful for preventing
PT and treating infectious diseases e.g. malaria and chronic disease e.g.
PT cancer, comprises peptide and carrier from protein D of influenzae -
XX
XX Disclosure; Page 7; 53pp; English.
XX
XX The present invention describes an immunogen (I) comprising a peptide
CC (Ia) and a carrier (Ib) derived from protein D of Haemophilus influenzae
CC or its fragment. Also described are: (1) a vaccine comprising (I), and
CC an excipient; (2) preparation of (I), comprising conjugating a peptide
CC to protein D or its fragment; and (3) preparation of a vaccine of (1),
CC comprising formulating (I) with an excipient. (I) has cytostatic,
CC antiallergic, nontropic, neuroprotective and protozoacide activities.
CC (1) and the vaccine are useful for the manufacture of a medicament for
CC preventing and treating infectious diseases such as malaria or chronic
CC disease such as cancer, Alzheimer's disease or allergy in a patient.
CC Unlike prior art immunogens, (I) induces high levels of antipeptide
CC immune responses while inducing a moderate humoral response against the
CC carrier. The present sequence represents an example of an immunogen from
CC the present invention which contains gonadotropin releasing hormone
CC (GnRH) tandem repeats.
XX
XX Sequence 23 AA;
XX
XX Query Match 37.9%; Score 92.5; DB 21; Length 23;
XX Best Local Similarity 43.2%; Pred. No. 3.2e-05;
XX Matches 19; Conservative 0; Mismatches 2; Indels 23; Gaps
XX
XX QY 2 HWSYGLRPGSSGSLQYIYKANSKFICITELSSGPSLHWSYGLRP 45
XX | | | | | | | | | | | | | | | | | | | | |
XX Db 2 HWSYGLRPGSC-----SEHWSYGLRP 22
XX
XX RESULT 4
XX AAR07323
XX ID AAR07323 standard; peptide; 30 AA.
XX
XX AAR07323;
XX
XX 29-JAN-1991 (first entry)
XX
XX Luteinising hormone releasing hormone derived peptide.
XX
XX LHRH; vaccine; meat; pigs; cancer; sterilisation.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 1 /label=OTHER
XX FT FT /note="OTHER=pyroglutamic acid or Gln having at
XX FT FT least one additional AA attached."
XX FT FT
XX Modified-site 3 /label=OTHER
XX FT FT /note="OTHER= N-formyl-Trp (optional)"
XX FT FT 13
XX Modified-site 13 /label=OTHER
XX FT FT /note="OTHER= N-formyl-Trp (optional)"
XX FT FT 10..19
XX Region /label=repeat
XX FT FT /note="-repeat must occur at least once"
XX FT FT 30
XX Misc-difference 30 /label=OTHER
XX FT FT /note="OTHER=Gly-NH2 or Gly having at
XX FT FT least one additional AA attached"
XX FT FT
XX WO9011298-A.
XX

```

[illegible]


```

XX 26-AUG-1999.
XX
XX 26-JAN-1999; 99WO-US01588.
XX
XX 19-FEB-1998; 98US-0026276.
XX
XX (IGEN-) IGEN INT INC.
XX
XX Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;
XX
XX WPI; 1999-518582/43.
XX
XX Epitope-containing fusion proteins used to generate a highly
XX specific immune responses
XX
XX Claim 81; Page 43; 67pp; English.
XX
XX This invention describes a novel fusion protein, comprising a heat shock
XX protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
XX which is useful for the stimulation of a highly specific immune response
XX when administered to an animal. The protein of the invention may be
XX post-translationally modified (e.g. by the addition of fatty acids to
XX enhance immunogenicity). The fusion proteins of the invention can be
XX used as vaccines to induce an immune response. When a T cell epitope is
XX attached, they can be used for control of viral infections, bacterial
XX infections, parasitic infection and cancer. The fusion proteins can be
XX used in pharmaceutical compositions for the treatment of gastrointestinal
XX diseases, pulmonary infections, respiratory infections, and HIV
XX infections. The use of ubiquitin as a scaffold is also useful for the
XX presentation and stimulation of anti-self immune responses, e.g.
XX generation of anti-gonadotropin releasing hormone antibodies which result
XX in the suppression of luteinizing hormone and follicle stimulating
XX hormone. This indirectly suppresses steroidogenesis and gamete maturation
XX in males and females. This type of anti-self response in humans is useful
XX in the treatment of prostate cancer and breast cancer. In livestock, the
XX ability to stimulate an anti-self response provides a simple alternative
XX to physical castration. Immunocastration of pigs is a better alternative
XX to physical castration, as it does not result in any of the detrimental
XX side effects associated with physical castration. Other examples of
XX diseases and conditions treated with self proteins fused with ubiquitin
XX are TWE and its epitopes to modulate septic shock, arthritis,
XX inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
XX epsilon heavy chain for the control of allergic reactions; chorionic
XX gonadotropin for fertility control; and sperm proteins for fertility
XX control. A further use of the fusion proteins is as part of a vaccine to
XX enhance growth rate and thereby the final weight of the livestock prior
XX to shipment to market. In addition, the fusion proteins of the invention
XX can be used to detect and identify antibodies from experimental samples.
XX This sequence represents a GnRH fragment used in the construction of
XX a ubiquitin fusion protein described in the method of the invention.
XX
XX SQ Sequence 41 AA;
    Query Match 37.3%; Score 91; DB 20; Length 41;
    Best Local Similarity 43.5%; Pred. No. 0.0001;
    Matches 20; Conservative 0; Mismatches 6; Indels 20; Gaps 2;
QY 2 HWSYGLRPGSGPSLQYIKANSKFTIGITELSSG--PSLHWSYGLR 45
    |||||
    | | | | |
Db 2 HWSYGLRPGQ-----HWSYGLRPGQHWSYGLR 29

RESULT 8
AAB71953
ID AAB71953 standard; Protein; 41 AA.
XX
XX AC AAB71953;
XX
XX 10-MAY-2001 (first entry)
XX
XX GnRH sequence #1.
XX

GnRH; gonadotropin releasing hormone; immunostimulant; antiallergic;
growth promoting; vaccine; heat shock fusion protein; ubiquitin,
self-epitope; immunogen; male-specific peptide hormone;
female-specific peptide hormone.
Unidentified.
WO2000112216-A1.
22-FEB-2001.
14-AUG-2000; 2000WO-US22121.
13-AUG-1999; 99US-0374721.
(PROT-) PROTEINIX CO.
13-AUG-1999; 99US-0374721.
(PROT-) PROTEINIX CO.
Kenten JH, Roberts S, Lohnas G;
WPI; 2001-202829/20.
Novel fusion proteins for stimulating immune response in animals
against self-antigen, has heat shock protein fused to single or a group
of epitope-containing segments having identical or non-identical self
epitopes -
Example 5; Page 44; 94pp; English.
The present sequence is given in an example illustrating an invention
disclosing self-epitope-containing heat shock fusion proteins. The
proteins comprise a heat shock protein, preferably ubiquitin, fused to a
single epitope-containing segment or two or more non-contiguous epitope
containing segments, each epitope-containing segment comprising one or
more identical or non-identical self-epitopes. The heat shock fusion
proteins are useful for stimulating an immune response in an animal,
especially a pig, directed towards a self-antigen, in particular a male
or female-specific peptide hormone. The physiological consequence is
substantially similar to the consequences of surgical castration. The
proteins are useful for identifying antibodies in experimental or
diagnostic samples and for reducing levels of a predetermined endogenous
protein, such as gonadotropin releasing hormone, tumour necrosis factor
or growth hormone protein. They are also useful for reducing allergic
response and for increasing the growth rate of an animal. The proteins
have a high specific immune response and do not result in detrimental
side-effects. The present sequence was used to create ubiquitin-GnRH
fusion proteins.
XX
XX SQ Sequence 41 AA;
    Query Match 37.3%; Score 91; DB 22; Length 41;
    Best Local Similarity 43.5%; Pred. No. 0.0001;
    Matches 20; Conservative 0; Mismatches 6; Indels 20; Gaps 2;
QY 2 HWSYGLRPGSGPSLQYIKANSKFTIGITELSSG--PSLHWSYGLR 45
    |||||
    | | | | |
Db 2 HWSYGLRPGQ-----HWSYGLRPGQHWSYGLR 29

RESULT 9
AAY96085
ID AAY96085 standard; Protein; 40 AA.
XX
XX AC AAY96085;
XX
XX 19-DEC-2000 (first entry)
XX
XX Cattle gonadotropin releasing hormone tetramer.
XX
XX Gonadotropin releasing hormone; GnRH; cattle; vaccine;

```

Antigenic peptide

26-AUG-1999

PF 26-JAN-1999; 99WO-US01588.
 XX 19-FEB-1998; 98US-0026276.
 XX (IGEN-) IGEN INT INC.
 XX Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;
 XX WPI; 1999-518582/43.
 DR Epitope-containing fusion proteins used to generate a highly
 XX specific immune responses
 XX Example 3; Page 40; 67pp; English.
 XX This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC presentation and stimulation of anti-self immune responses, e.g.
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steroidogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful
 CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TNF and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC epsilon heavy chain for the control of allergic reactions; chorionic
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion proteins is as part of a vaccine to
 CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH dimer used in the construction of a
 CC ubiquitin fusion protein described in the method of the invention.
 XX Sequence 20 AA;
 SQ Query Match 35.7%; Score 87; DB 20; Length 20;
 Best Local Similarity 38.6%; Pred. No. 0.00014;
 Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;
 QY 2 HWSYGLRPGSGPSLQYIKANSKFCITELSSGPSLHWSYGLRP 45
 DB 2 HWSYGLRPGQ-----HWSYGLRP 19
 RESULT 12
 AAY31177
 ID AAY31177 standard; peptide: 20 AA.
 XX AC AAY31177;
 XX 28-OCT-1999 (first entry)
 XX Ubiquitin fusion protein GnRH dimer.
 DE Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;

KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steroidogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.
 XX Synthetic.
 OS WO9942472-A1.
 XX PN 26-AUG-1999.
 XX PD 26-JAN-1999; 99WO-US01588.
 XX PF 19-FEB-1998; 98US-0026276.
 XX PR (IGEN-) IGEN INT INC.
 XX PA Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;
 XX PI WPI; 1999-518582/43.
 XX DR Epitope-containing fusion proteins used to generate a highly
 XX specific immune responses
 XX Example 3; Page 41; 67pp; English.
 XX This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC presentation and stimulation of anti-self immune responses, e.g.
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steroidogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful
 CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TNF and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC epsilon heavy chain for the control of allergic reactions; chorionic
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion proteins is as part of a vaccine to
 CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH dimer used in the construction of
 CC a ubiquitin fusion protein described in the method of the invention.
 XX Sequence 20 AA;
 SQ Query Match 35.7%; Score 87; DB 20; Length 20;
 Best Local Similarity 38.6%; Pred. No. 0.00014;
 Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

QY 2 HWSYGLRPGSGPSLQYIKANSKFCITELSSGPSLHWSYGLRP 45
 DB 2 HWSYGLRPGQ-----HWSYGLRP 19
 RESULT 13

AA311178	Sequence	20 AA;	Query Match	35.7%;	Score 87;	DB 20;	Length 20;
ID	AA311178 standard; peptide; 20 AA.		Best Local Similarity	38.6%;	Pred. No. 0.00014;		
AC	AA311178;		Matches	17;	Conservative	0;	Mismatches
DT	28-OCT-1999 (first entry)					1;	Indels
XX	Ubiquitin fusion protein GnRH mixed dimer 1.					26;	Gaps
XX							
DE	Ubiquitin fusion protein GnRH mixed dimer 1.						
XX							
KW	Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;						
KW	immune response stimulation; vaccine; T cell; viral; infection; cancer;						
KW	bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;						
KW	pulmonary infection; respiratory infection; scaffold; anti-self; pig;						
KW	steroidogenesis; gamete maturation; prostate; breast; castration; TNF;						
KW	tumour necrosis factor; septic shock; arthritis; Crohn's disease;						
KW	inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;						
KW	fertility; sperm protein; growth rate; antibody; detection; GnRH.						
XX							
OS	Synthetic.						
XX							
PN	WO9942472-A1.						
XX							
PD	26-AUG-1999.						
XX							
PF	26-JAN-1999; 99WO-US01588.						
XX							
PR	19-FEB-1998; 98US-0026276.						
XX							
PA	(IGEN-) IGEN INT INC.						
XX							
PI	Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;						
XX							
DR	WPI; 1999-518582/43.						
XX							
PT	Epitope-containing fusion proteins used to generate a highly						
PT	specific immune responses						
XX							
PS	Example 3; Page 41; 67pp; English.						
XX							
CC	This invention describes a novel fusion protein, comprising a heat shock						
CC	protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner						
CC	which is useful for the stimulation of a highly specific immune response						
CC	when administered to an animal. The protein of the invention may be						
CC	post-translationally modified (e.g. by the addition of fatty acids to						
CC	enhance immunogenicity). The fusion proteins of the invention can be						
CC	used as vaccines to induce an immune response. When a T cell epitope is						
CC	attached, they can be used for control of viral infections, bacterial						
CC	infections, parasitic infection and cancer. The fusion proteins can be						
CC	used in pharmaceutical compositions for the treatment of gastrointestinal						
CC	diseases, pulmonary infections, respiratory infections, and HIV						
CC	infections. The use of ubiquitin as a scaffold is also useful for the						
CC	presentation and stimulation of anti-self immune responses, e.g.						
CC	generation of anti-gonadotropin releasing hormone antibodies which result						
CC	in the suppression of luteinizing hormone and follicle stimulating						
CC	hormone. This indirectly suppresses steroidogenesis and gamete maturation						
CC	in males and females. This type of anti-self response in humans is useful						
CC	in the treatment of prostate cancer and breast cancer. In livestock, the						
CC	ability to stimulate an anti-self response provides a simple alternative						
CC	to physical castration. Immunocastration of pigs is a better alternative						
CC	to physical castration, as it does not result in any of the detrimental						
CC	side effects associated with physical castration. Other examples of						
CC	diseases and conditions treated with self proteins fused with ubiquitin						
CC	are TNF and its epitopes to modulate septic shock, arthritis,						
CC	inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig						
CC	epsilon heavy chain for the control of allergic reactions; chorionic						
CC	gonadotropin for fertility control; and sperm proteins for fertility						
CC	control. A further use of the fusion proteins is as part of a vaccine to						
CC	enhance growth rate and thereby the final weight of the livestock prior						
CC	to shipment to market. In addition, the fusion proteins of the invention						
CC	can be used to detect and identify antibodies from experimental samples.						
CC	This sequence represents a GnRH mixed dimer used in the construction of						
CC	a ubiquitin fusion protein described in the method of the invention.						
XX							

CC diseases and conditions treated with self proteins fused with ubiquitin
CC are TNF and its epitopes to modulate septic shock, arthritis,
CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
CC epsilon heavy chain for the control of allergic reactions; chorionic
CC gonadotropin for fertility control; and sperm proteins for fertility
CC control. A further use of the fusion proteins is as part of a vaccine to
CC enhance growth rate and thereby the final weight of the livestock prior
CC to shipment to market. In addition, the fusion proteins of the invention
CC can be used to detect and identify antibodies from experimental samples.
CC This sequence represents a GnRH mixed dimer used in the construction of
CC a ubiquitin fusion protein described in the method of the invention.
XX

SQ Sequence 20 AA;

Query Match 35.7%; Score 87; DB 20; Length 20;
Best Local Similarity 38.6%; Pred. NO. 0.00014;
Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

QY 2 HWSYGLRPGSSGPSLQYIKANSKFGITELSSGPSLHWSYGLRP 45
||||||| |
Db 2 HWSYGLRPGEE-----HWSYGLRP 19

RESULT 15

AAB71945
ID AAB71945 standard; Peptide: 20 AA.

XX AC AAB71945;

XX DT 10-MAY-2001 (first entry)

XX DE GnRH dimer.

XX KW GnRH; gonadotropin releasing hormone; immunostimulant; antiallergic;
XX growth promoting; vaccine; heat shock fusion protein; ubiquitin;
XX self-epitope; immunogen; male-specific peptide hormone;
XX female-specific peptide hormone.

XX OS Unidentified.

XX PN WO200112216-A1.

XX PD 22-FEB-2001.

XX PF 14-AUG-2000; 2000WO-US22121.

XX PR 13-AUG-1999; 99US-0374721.

XX PA (PROT-) PROTEINIX CO.

XX PR 13-AUG-1999; 99US-0374721.

XX PA (PROT-) PROTEINIX CO.

XX PI Kenten JH, Roberts S, Lohnas G;

XX DR WPI; 2001-202829/20.

XX Novel fusion proteins for stimulating immune response in animals
PT against self-antigen, has heat shock protein fused to single or a group
PT of epitope-containing segments having identical or non-identical self
PT epitopes -
XX

PS Claim 14; Page 63; 94pp; English.

XX The present sequence is given in an example illustrating an invention
CC disclosing self-epitope-containing heat shock fusion proteins. The
CC proteins comprise a heat shock protein, preferably ubiquitin, fused to a
CC single epitope-containing segment or two or more non-contiguous epitope
CC containing segments, each epitope-containing segment comprising one or
CC more identical or non-identical self-epitopes. The heat shock fusion
CC proteins are useful for stimulating an immune response in an animal,
CC especially a pig, directed towards a self-antigen, in particular a male

CC or female-specific peptide hormone. The physiological consequence is
CC substantially similar to the consequences of surgical castration. The
CC proteins are useful for identifying antibodies in experimental or
CC diagnostic samples and for reducing levels of a predetermined endogenous
CC protein, such as gonadotropin releasing hormone, tumour necrosis factor
CC or growth hormone protein. They are also useful for reducing allergic
CC response and for increasing the growth rate of an animal. The proteins
CC have a high specific immune response and do not result in detrimental
CC side-effects.

SQ Sequence 20 AA;

Query Match 35.7%; Score 87; DB 22; Length 20;
Best Local Similarity 38.6%; Pred. NO. 0.00014;
Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

QY 2 HWSYGLRPGSSGPSLQYIKANSKFGITELSSGPSLHWSYGLRP 45
||||||| |

Db 2 HWSYGLRPGQ-----HWSYGLRP 19

Search completed: October 10, 2002, 23:04:29
Job time : 28 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	113.5	46.5	27	1	US-08-446-692-13	Sequence 13, Appl	
2	113.5	46.5	27	1	US-08-488-351A-13	Sequence 13, Appl	
3	91	37.3	40	4	US-09-026-276-35	Sequence 35, Appl	
4	91	37.3	41	4	US-09-026-276-34	Sequence 34, Appl	
5	90	36.9	44	1	US-07-690-983D-45	Sequence 45, Appl	
6	87	35.7	20	4	US-07-690-983D-40	Sequence 40, Appl	
7	87	35.7	20	4	US-09-026-276-26	Sequence 26, Appl	
8	87	35.7	20	4	US-09-026-276-29	Sequence 29, Appl	
9	87	35.7	20	4	US-09-026-276-30	Sequence 30, Appl	
10	87	35.7	20	4	US-09-026-276-31	Sequence 31, Appl	
11	87	35.7	24	1	US-07-690-983D-43	Sequence 43, Appl	
12	84	34.4	24	5	PT- US92-07218-31	Sequence 31, Appl	
13	76	31.1	17	5	PT- US95-08596-23	Sequence 23, Appl	
14	76	31.1	31	5	PT- US93-11703-63	Sequence 63, Appl	
15	76	31.1	37	1	US-08-446-692-57	Sequence 57, Appl	
16	76	31.1	37	1	US-08-446-692-63	Sequence 63, Appl	
17	76	31.1	37	2	US-08-488-351A-57	Sequence 57, Appl	
18	76	31.1	37	2	US-08-488-351A-63	Sequence 63, Appl	
19	74	30.3	15	2	US-08-319-704-10	Sequence 10, Appl	
20	74	30.3	15	2	US-08-661-052-6	Sequence 6, Appl	
21	74	30.3	15	2	US-08-460-503-7	Sequence 7, Appl	
22	74	30.3	15	4	US-09-046-373-2	Sequence 2, Appl	
23	74	30.3	15	4	PT- US93-11703-69	Sequence 69, Appl	
24	74	30.3	15	5	PT- US98-188-083-6	Sequence 6, Appl	
25	74	30.3	16	4	US-09-248-588-55	Sequence 55, Appl	
26	74	30.3	17	1	US-08-446-692-4	Sequence 4, Appl	
27	74	30.3	17	2	US-08-488-351A-4	Sequence 4, Appl	

```
RESULT 2
US-08-488-351A-13
; Sequence 13, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang YI
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-13

Query Match 46.5%; Score 113.5; DB 2; Length 27;
Best Local Similarity 79.3%; Pred. No. 3.1e-08;
Matches 23; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 17 QYIKANSKFIGITELSSGPSLHWSYGLRP 45
Db 3 QYIKANSKFIGITELE-----HWSYGLRP 26

RESULT 3
US-09-026-276-35
; Sequence 35, Application US/09026276
; Patent No. 6319503
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, Aprille L
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; CURRENT APPLICATION NUMBER: US/09/026,276
; CURRENT FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Porcine
US-09-026-276-35

Query Match 37.3%; Score 91; DB 4; Length 40;
Best Local Similarity 43.5%; Pred. No. 3.6e-05;
Matches 20; Conservative 0; Mismatches 6; Indels 20; Gaps 2;

Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSG--PSLHWSYGLRP 45
Db 2 HWSYGLRPGQ-----HWSYGLRPGQHWSYGLRP 29

RESULT 4
US-09-026-276-34
; Sequence 34, Application US/09026276
; Patent No. 6319503
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, Aprille L
; APPLICANT: Lohnas, Gerald L
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; CURRENT APPLICATION NUMBER: US/09/026,276
; CURRENT FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Porcine
US-09-026-276-34

Query Match 37.3%; Score 91; DB 4; Length 41;
Best Local Similarity 43.5%; Pred. No. 3.7e-05;
Matches 20; Conservative 0; Mismatches 6; Indels 20; Gaps 2;

Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSG--PSLHWSYGLRP 45
Db 2 HWSYGLRPGQ-----HWSYGLRPGQHWSYGLRP 29

RESULT 5
US-07-690-983D-45
; Sequence 45, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-690-983D-45

Query Match 36.9%; Score 90; DB 1; Length 44;
Best Local Similarity 43.5%; Pred. No. 5.4e-05;
Matches 20; Conservative 0; Mismatches 6; Indels 20; Gaps 2;

Qy 2 HWSYGLRPGSSGSPSLQYIKANSKFIGITELSSG--PSLHWSYGLRP 45
Db 4 HWSYGLRPG-----HWSYGLRPGEHWSYGLRP 31
|||||
| | | | |

RESULT 6
US-07-690-983D-40
; Sequence 40, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
US-07-690-983D-40

Query Match 35.7%; Score 87; DB 1; Length 20;
Best Local Similarity 38.6%; Pred. No. 5.2e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

Qy 2 HWSYGLRPGSSGSPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 3 HWSYGLRPG-----HWSYGLRP 20
|||||
| | | | |

RESULT 7
US-09-026-276-26
; Sequence 26, Application US/09026276
; Patent No. 6319503
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, Aprille L
; APPLICANT: Lohnas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: U.S. Patent Application No. 6319503 09\026,276
; CURRENT APPLICATION NUMBER: US/09/026,276
; CURRENT FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Porcine
; US-09-026-276-26

Query Match 35.7%; Score 87; DB 4; Length 20;
Best Local Similarity 38.6%; Pred. No. 5.2e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

Qy 2 HWSYGLRPGSSGSPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 2 HWSYGLRPGQ-----HWSYGLRP 19
|||||
| | | | |

RESULT 8
US-09-026-276-29
; Sequence 29, Application US/09026276
; Patent No. 6319503
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, Aprille L
; APPLICANT: Lohnas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: U.S. Patent Application No. 6319503 09\026,276
; CURRENT APPLICATION NUMBER: US/09/026,276
; CURRENT FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Porcine
; US-09-026-276-29

Query Match 35.7%; Score 87; DB 4; Length 20;
Best Local Similarity 38.6%; Pred. No. 5.2e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

Qy 2 HWSYGLRPGSSGSPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 2 HWSYGLRPG-----HWSYGLRP 19
|||||
| | | | |
```

RESULT 9
US-09-026-276-30
; Sequence 30, Application US/09026276
; Patent No. 6319503
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, Aprille L
; APPLICANT: Lohnas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: U.S. Patent Application No. 6319503 09\026,276
; CURRENT APPLICATION NUMBER: US/09\026,276
; CURRENT FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Porcine
US-09-026-276-30

Query Match 35.7%; Score 87; DB 4; Length 20;
Best Local Similarity 38.6%; Pred. No. 5.2e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

Qy 2 HWSYGLRPGSSGSLQYIKANSKFITELSSGSPSLHWSYGLRP 45
|||||
Db 2 HWSYGLRPGQ-----HWSYGLRP 19
|||||

RESULT 10
US-09-026-276-31
; Sequence 31, Application US/09026276
; Patent No. 6319503
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, Aprille L
; APPLICANT: Lohnas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: U.S. Patent Application No. 6319503 09\026,276
; CURRENT APPLICATION NUMBER: US/09\026,276
; CURRENT FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Porcine
US-09-026-276-31

Query Match 35.7%; Score 87; DB 4; Length 20;
Best Local Similarity 38.6%; Pred. No. 5.2e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

Qy 2 HWSYGLRPGSSGSLQYIKANSKFITELSSGSPSLHWSYGLRP 45
|||||
Db 2 HWSYGLRPGQ-----HWSYGLRP 19
|||||

RESULT 11
US-07-690-983D-43
; Sequence 43, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/690,983D
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-690-983D-43

Query Match 35.7%; Score 87; DB 1; Length 24;
Best Local Similarity 38.6%; Pred. No. 6.5e-05;
Matches 17; Conservative 0; Mismatches 1; Indels 26; Gaps 1;

Qy 2 HWSYGLRPGSSGSLQYIKANSKFITELSSGSPSLHWSYGLRP 45
|||||
Db 4 HWSYGLRPGQ-----HWSYGLRP 21
|||||

RESULT 12
PCT-US92-07218-31
; Sequence 31, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/07218
; FILING DATE: 19920826
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14137-26-3
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2400
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-07218-31

Query Match 34.4%; Score 84; DB 5; Length 24;
Best Local Similarity 77.3%; Pred. No. 0.00016;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 PGSSGPSLQYIKANSKFIGITE 30
Db 3 PSDFPFSVQYIKANSKFIGITE 24

RESULT 13
PCT-US95-08596-23
Sequence 23, Application PC/TUS9508596
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Proinsulin Peptide Compounds for Detecting
and Treating Type I Diabetes
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08596
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/272,220
FILING DATE: 08-JULY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: DGI-092PC
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
PCT-US95-08596-23

Query Match 31.1%; Score 76; DB 5; Length 17;
Best Local Similarity 93.8%; Pred. No. 0.0011;

Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 16 LOYIKANSKFIGITEL 31
Db 2 MOYIKANSKFIGITEL 17

RESULT 14
PCT-US93-11703-63
Sequence 63, Application PC/TUS9311703
GENERAL INFORMATION:
APPLICANT: Chiron Mimotopes Pty. Ltd.
TITLE OF INVENTION: T-Cell Epitopes
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grant D. Green
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11703
FILING DATE: 28-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,852
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0222.101
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-11703-63

Query Match 31.1%; Score 76; DB 5; Length 31;
Best Local Similarity 93.8%; Pred. No. 0.0022;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LOYIKANSKFIGITEL 31
Db 9 MOYIKANSKFIGITEL 24

RESULT 15
US-08-446-692-57
Sequence 57, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Y1
APPLICANT: Zamb, Timotby
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York

; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-57

Query Match 31.1%; Score 76; DB 1; Length 37;
Best Local Similarity 84.2%; Pred. No. 0.0027;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 13 GPSLOYIKANSKFIGITEL 31
| | | | | | | | | | | | | | | | | | | | |
Db 17 GKKQYIKANSKFIGITEL 35

Search completed: October 10, 2002, 23:07:04
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 23:08:06 ; Search time 16 Seconds
(without alignments)
306.285 Million cell updates/sec

Title: US-09-848-834a-20

Perfect score: 265

Sequence: 1 XHSYGLRPGSGPSLDEKK.....NVNSSGSPSLHSYGLRPGX 51

Scoring table: BLOSUMP62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 12171

Minimum DB seq length: 0

Maximum DB seq length: 51

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	21.9	10	1	RHPGG
2	58	21.9	10	1	RHSHG
3	54	20.4	10	1	RHAQ1
4	45	17.0	10	2	A21114
5	40	15.1	10	1	RHAQ2
6	40	15.1	10	1	A61126
7	40	15.1	10	2	A46030
8	40	15.1	10	2	B46030
9	40	15.1	30	2	F70254
10	38	14.3	44	2	S28770
11	37.5	14.2	32	2	S20721
12	37.5	14.2	42	2	C45095
13	37.5	14.2	50	2	G83624
14	37	14.0	10	2	A49187
15	37	14.0	27	2	I54247
16	37	14.0	35	2	A59371
17	37	14.0	44	2	C24691
18	36.5	13.8	47	2	T37586
19	36	13.6	35	2	B85708
20	35.5	13.4	39	2	S50362
21	35.5	13.4	49	2	S65722
22	35	13.2	22	2	B40256
23	35	13.2	32	2	A61624
24	35	13.2	38	2	C82373
25	35	13.2	48	2	G81543
26	35	13.2	49	2	G69538
27	35	13.2	51	2	E72807
28	34.5	13.0	51	2	S36377
29	34	12.8	39	2	H86068

30	34	12.8	44	2	A82079
31	34	12.8	45	2	E83937
32	34	12.8	48	2	A46232
33	34	12.8	49	2	S72213
34	33.5	12.6	40	2	S33393
35	33.5	12.6	42	2	B48301
36	33.5	12.6	48	2	A49007
37	33.5	12.6	50	2	F71281
38	33	12.5	16	2	S7517
39	33	12.5	32	2	A56634
40	33	12.5	48	2	AC1020
41	33	12.5	49	2	H90606
42	32.5	12.3	18	2	A40256
43	32.5	12.3	21	2	C60275
44	32.5	12.3	41	2	A26688
45	32	12.1	17	4	I51887

ALIGNMENTS

RESULT 1

RHPGG
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A:Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 21.9%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
| | | | | | | |
Db 2 HWSYGLRPG 10

RESULT 2

RHSHG
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780
A:Molecule type: protein

A:Residues: 1-10 <BUR>

A:Note: the natural and synthetic hormones have the same biological activity
 C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormones
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 21.9%; Score 58; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.51;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10

RESULT 3

RHAQ1

gonadoliberin I - American alligator
 N:Alternate names: gonadotropin-releasing hormone I
 C:Species: Alligator mississippiensis (American alligator)
 C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
 C:Accession: A60066
 R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson
 Regul. Pept. 33, 105-116, 1991
 A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
 A:Reference number: A60066; MUID:91352338

A:Accession: A60066

A:Molecule type: protein

A:Residues: 1-10 <LOW>

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 20.4%; Score 54; DB 1; Length 10;

Best Local Similarity 88.9%; Pred. No. 1.5;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10

Db 2 HWSYGLQPG 10

RESULT 4

A21114

gonadoliberin - chum salmon

C:Species: Oncorhynchus keta (chum salmon)

C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993

C:Accession: A21114

R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.

Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983

A:Title: Characterization of a teleost gonadotropin-releasing hormone.

A:Reference number: A21114; MUID:83195140

A:Accession: A21114

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <SHE>

Query Match 17.0%; Score 45; DB 2; Length 10;

Best Local Similarity 77.8%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10

Db 2 HWSYGLWLP 10

RESULT 5

RHAQ2

gonadoliberin·II - American alligator

N:Alternate names: gonadotropin-releasing hormone II

C:Species: Alligator mississippiensis (American alligator)

C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997

C:Accession: B60066

R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan

Regul. Pept. 33, 105-116, 1991

A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator

A:Reference number: A60066; MUID:91352338

A:Accession: B60066

A:Molecule type: protein

A:Residues: 1-10 <LOW>

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 15.1%; Score 40; DB 1; Length 10;

Best Local Similarity 66.7%; Pred. No. 74;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10

Db 2 HWSHGWP 10

RESULT 6

A61126

gonadoliberin - spotted ratfish

N:Alternate names: gonadotropin-releasing hormone

C:Species: Hydrotaea colliei (spotted ratfish)

C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997

C:Accession: A61126

R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.

Gen. Comp. Endocrinol. 82, 152-161, 1991

A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a holocarpine ratfish

A:Reference number: A61126; MUID:91340067

A:Accession: A61126

A:Molecule type: protein

A:Residues: 1-10 <LOW>

A:Experimental source: brain

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 15.1%; Score 40; DB 1; Length 10;

Best Local Similarity 66.7%; Pred. No. 74;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10

Db 2 HWSHGWP 10

RESULT 7

A46030

gonadoliberin I - spiny dogfish

N:Alternate names: gonadotropin-releasing hormone

C:Species: Squalus acanthias (spiny dogfish)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998

C:Accession: A46030

R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Pete

Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992

A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain

A:Reference number: A46030; MUID:92335300

A:Accession: A46030

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <LOW>

C:Keywords: hormone; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 15.1%; Score 40; DB 2; Length 10;

Best Local Similarity 66.7%; Pred. No. 74;

Matches 36; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWLP 10

RESULT 8
B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squallus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, F.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 15.1%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWYP 10

RESULT 9
F70254
hypothetical protein BBK29 - Lyme disease spirochete plasmid K/lp36
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: F70254
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: F70254
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-30 <KLE>
A:Cross-references: GB:AE000788; NID:g2690123; PIDN:AAG66163.1; PID:g2690155; TIGR:BBK29
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 15.1%; Score 40; DB 2; Length 30;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 8; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 10 GSSGSLDEKTKAKMEKASSVFN 33
DB 2 GNKPSSGDKMKIRKLSSYYKI 25

RESULT 10
S28770
hypothetical protein 44 (psbI 3' region) - barley chloroplast
C:Species: chloroplast Hordeum vulgare (barley)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C:Accession: S28770
R:Sexton, T.B.; Jones, J.T.; Mullet, J.E.
Curr. Genet. 17, 445-454, 1990

Matches 36; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWLP 10

RESULT 8
B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squallus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, F.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 15.1%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWYP 10

RESULT 9
F70254
hypothetical protein BBK29 - Lyme disease spirochete plasmid K/lp36
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: F70254
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: F70254
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-30 <KLE>
A:Cross-references: GB:AE000788; NID:g2690123; PIDN:AAG66163.1; PID:g2690155; TIGR:BBK29
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 15.1%; Score 40; DB 2; Length 30;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 8; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 10 GSSGSLDEKTKAKMEKASSVFN 33
DB 2 GNKPSSGDKMKIRKLSSYYKI 25

RESULT 10
S28770
hypothetical protein 44 (psbI 3' region) - barley chloroplast
C:Species: chloroplast Hordeum vulgare (barley)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C:Accession: S28770
R:Sexton, T.B.; Jones, J.T.; Mullet, J.E.
Curr. Genet. 17, 445-454, 1990

Matches 36; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWLP 10

RESULT 8
B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squallus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, F.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 15.1%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWYP 10

RESULT 9
F70254
hypothetical protein BBK29 - Lyme disease spirochete plasmid K/lp36
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: F70254
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kervatage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: F70254
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-30 <KLE>
A:Cross-references: GB:AE000788; NID:g2690123; PIDN:AAG66163.1; PID:g2690155; TIGR:BBK29
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 15.1%; Score 40; DB 2; Length 30;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 8; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 10 GSSGSLDEKTKAKMEKASSVFN 33
DB 2 GNKPSSGDKMKIRKLSSYYKI 25

RESULT 10
S28770
hypothetical protein 44 (psbI 3' region) - barley chloroplast
C:Species: chloroplast Hordeum vulgare (barley)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C:Accession: S28770
R:Sexton, T.B.; Jones, J.T.; Mullet, J.E.
Curr. Genet. 17, 445-454, 1990

A:Title: Sequence and transcriptional analysis of the barley ctDNA region upstream of
A:Reference number: S28765; MUID:90291518
A:Accession: S28770
A:Molecule type: DNA
A:Residues: 1-44 <SEX>
A:Cross-references: EMBL:X52765; NID:g11601; PIDN:CAA36977.1; PID:g11607
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 14.3%; Score 38; DB 2; Length 44;
Best Local Similarity 43.5%; Pred. No. 6.8e+02;
Matches 10; Conservative 1; Mismatches 8; Indels 4; Gaps 1;

QY 3 WSYGLRPGSSPSLDEKTKAKME 25
DB 19 WRPGFRPG-----SLDKNPKIKSE 37

RESULT 11
S20771
Ig heavy chain V region (VH4, INLP24) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 23-Jul-1999
C:Accession: S20771
R:Mortari, F.; Wang, J.; Schroeder, H.W.
submitted to the EMBL Data Library, April 1992
A:Description: Analysis of human cord blood Ig heavy chain IgA and IgG repertoire.
A:Reference number: S20764
A:Accession: S20771
A:Molecule type: DNA
A:Residues: 1-32 <MOR>
A:Cross-references: EMBL:Z11955; NID:g33871; PIDN:CAA78012.1; PID:g33872
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match 14.2%; Score 37.5; DB 2; Length 32;
Best Local Similarity 40.0%; Pred. No. 5.5e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 28 SSVFNVNVSNGSPSLHWSYG 47
DB 5 TAVYYCTREYSGPSL-WDWG 23

RESULT 12
C45095
photosystem I light-harvesting complex chlorophyll a/b protein, p14.1 - Chlamydomonas
C:Species: Chlamydomonas reinhardtii
C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 21-Aug-1998
C:Accession: C45095
R:Bassi, R.; Soen, S.Y.; Frank, G.; Zuber, H.; Rochaix, J.D.
J. Biol. Chem. 267, 25714-25721, 1992
A:Title: Characterization of chlorophyll a/b proteins of photosystem I from Chlamydom
A:Reference number: A45095; MUID:93100280
A:Accession: C45095
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-42 <BAS>
A>Note: sequence extracted from NCBI backbone (NCBIP:120935)
C:Superfamily: chlorophyll a/b-binding protein

Query Match 14.2%; Score 37.5; DB 2; Length 42;
Best Local Similarity 30.8%; Pred. No. 7.4e+02;
Matches 12; Conservative 6; Mismatches 16; Indels 5; Gaps 1;

QY 17 DEKTKAKMEKASSVFNVSNG-----PSLHWSYGLRP 50
DB 1 EEKSIKAKVDRSKDQLAVGASQSSLAYLDGSLGDFGDP 39

RESULT 13
G83624

hypothetical protein PA0161 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83624
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: G83624
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-50 <STO>
A:Cross-references: GB:AB004454; GB:AE004091; NID:g9945990; PIDN:AAG03551.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0161

Query Match 14.2%; Score 37.5; DB 2; Length 50;
Best Local Similarity 32.3%; Pred. No. 9e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 9; Gaps 2;

QY 17 DEKTIKMEKASSVNVNSSG---PSLHW 44
||| :||: | | | | | | | |
DB 7 DESKRLSIEESDCV-----SAGFQCPEIHW 31

RESULT 14
A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A:Title: Primary structure and biological activity of a third gonadotropin-releasing hor
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 14.0%; Score 37; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
||| :||
DB 2 HWSHDWKEG 10

RESULT 15
I54247
GTP-binding regulatory protein Gt alpha-2 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000
C:Accession: I54247
R:Hirano, T.
Hokkaido Igaku Zasshi 68, 885-893, 1993
A:Title: [Analysis of cell specific transcription of the human cone transducin alpha sub
A:Reference number: I54247; MUID:94156312
A:Accession: I54247
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-27 <RES>
A:Cross-references: GB:S70567; NID:g545549; PIDN:AAD14063.1; PID:g4261763
C:Superfamily: GTP-binding regulatory protein Gs alpha chain

Query Match 14.0%; Score 37; DB 2; Length 27;
Best Local Similarity 46.7%; Pred. No. 5.2e+02;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 12 SGPSLDEKKIAKMEK 26
||| :||: | | | | | | | |
DB 3 SGASAEDEKELAKRSK 17

Search completed: October 10, 2002, 23:12:40
Job time : 17 secs

RESULT 2	
GONI_CLUPA	
ID GON1_CLUPA STANDARD; PRT; 10 AA.	
AC P81749;	
DT 16-OCT-2001 (Rel. 40, Created)	
DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DE 16-OCT-2001 (Rel. 40, Last annotation update)	
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH)	
DE (luliberin I).	
GN GNRH1.	
OS Clupea pallasii (Pacific herring).	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;	
Clupeinae; Clupeini; Clupeus; Clupeus	

OC Clupeinae: Clupea.
OX NCBI_TaxID=30724;
RN [1]
RP SEQUENCE, AND FUNCTION
RC TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR InterPro: IPR002012; GnRH.
DR PROSITE: PS00446; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA: 1105 MW; 284B20B72871F5A3 CRC64;

Query Match 17.7% Score 47; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.6; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 1;

Qy 2 HWSYGLRPG 10
| | | | | | | |
Db 2 HWSHGLSPG 10
| | | | | | | |

RESULT 3
GON3_ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GnRH-III) (LH-
RH III) (Luliberin III).
GN GnRH3.
OS Oncorhynchus keta (Chum salmon), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O. keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone."
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR InterPro: IPR002012; GnRH.
DR PROSITE: PS00446; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA: 1230 MW; 284B3233786B45A3 CRC64;

Query Match 17.0% Score 45; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 6.4; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2;

Qy 2 HWSYGLRPG 10
| | | | | | | |
Db 2 HWSYGLWLP 10
| | | | | | | |

RESULT 4
GON2_CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GnRH-II)
(LH-RH II) (Luliberin II).
OS Gallus gallus (Chicken).
OS Alligator mississippiensis (American alligator),
OS Squalus acanthias (Spiny dogfish), and
OS Hydrolagus colliei (Spotted ratfish) (Pacific ratfish), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=Chicken; TISSUE=Hypothalamus;
RX MEDLINE=84222059; PubMed=6427779;
RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
Matsuo H.;
RT "Identification of the second gonadotropin-releasing hormone in
RT chicken hypothalamus: evidence that gonadotropin secretion is
RT probably controlled by two distinct gonadotropin-releasing hormones
RT in avian species."
RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=A.mississippiensis; TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
RT from brains of the American alligator (Alligator mississippiensis)."
RL Regul. Pept. 33:105-116(1991).
RN [3]
RP SEQUENCE.
RC SPECIES=S. acanthias; TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
RT dogfish brain provides insight into GnRH evolution."
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN [4]
RP SEQUENCE.
RC SPECIES=H.colliei; TISSUE=Brain;
RX MEDLINE=91340067; PubMed=1678723;
RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
Lee T.;
RT "Primary structure of gonadotropin-releasing hormone from the brain
RT of a holoccephalan (ratfish: Hydrolagus colliei)."
RL Gen. Comp. Endocrinol. 82:152-161(1991).
RN [5]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

```

RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RT hormones, including a novel form, from an ancient teleost, herring.";
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: B60066; RHAQ2.
DR PIR: A61126; A61126.
DR PIR: B46030; B46030.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA: 1254 MW: 284B2E437871F5A3 CRC64;

Query Match 15.1%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
   |||: ||
Db 2 HWSHGWPY 10

RESULT 5
GONL_SQUAC STANDARD; PRT; 10 AA.
AC P27429;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
DE (Luliberin).
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidel; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=923335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
RT dogfish brain provides insight into GNRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA: 1204 MW: 284B32337871F5A3 CRC64;

Query Match 15.1%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
   |||: ||
Db 2 HWSHGWPY 10

RESULT 6
GON3_PETMA STANDARD; PRT; 10 AA.
AC P30948;

```

```

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
DE (Luliberin III).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Loves S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
RT releasing hormone from lamprey brain.";
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA: 1277 MW: 284B36237AA1F5A3 CRC64;

Query Match 14.0%; Score 37; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
   |||: ||
Db 2 HWSHDWPKG 10

RESULT 7
APL3_DIAGR STANDARD; PRT; 32 AA.
AC P81471;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Apolipophorin-III (APOLP-III) (Fragment).
OS Diatraea grandiosella (Southwestern corn borer).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Pyraloidea; Pyralidae; Crambinae; Diatraea.
OX NCBI_TaxID=61289;
RN [1]
RP SEQUENCE.
RC TISSUE=Larval plasma;
RA Burks C.S., Shelby K.S., Chippendale G.M.;
RT "Characteristics of apolipophorin-III of the southwestern corn borer,
RT Diatraea grandiosella.";
RL Insect Biochem. Mol. Biol. 22:905-915(1992).
CC -!- FUNCTION: INCREASE OF LIPID CARRYING CAPACITY OF LIPOPHORIN BY
CC COVERING THE EXPANDING HYDROPHOBIC SURFACE RESULTING FROM DIACYL-
CC GLYCEROL UPTAKE.
CC -!- SUBUNIT: EQUILIBRIUM BETWEEN A SOLUBLE MONOMER AND A BOUND
CC LIPOPROTEIN FORM. APOLIPOPHORIN-III ASSOCIATES WITH LIPOPHORIN
CC DURING LIPID LOADING UNTIL EACH PARTICLE CONTAINS 9 OR 14
CC MOLECULES OF APOLIPOPHORIN-III.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR; SECRETED (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: HEMOLYMPH.
CC -!- SIMILARITY: BELONGS TO THE INSECT APOLIPOPHORIN-III FAMILY.
KW Lipid transport; Hemolymph.
FT NON_TER 32 32
SQ SEQUENCE 32 AA: 3584 MW: A6C2D6D675C0B22 CRC64;

Query Match 13.2%; Score 35; DB 1; Length 32;

```

Best Local Similarity 38.9%; Pred. NO. 4e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 9 PGSSGSLDEKKIAKWEK 26
Db 3 PSTTPPDQXKKADEFQK 20

RESULT 8

VG62_BPMD2
ID VG62_BPMD2 STANDARD; PRT; 51 AA.
AC O64254;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gene 62 protein (GP62).
GN 62.
OS Mycobacteriophage D29.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=28369;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: implications for phage evolution.";
RL J. Mol. Biol. 279:143-164(1998).
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF022214; AAC18504.1; -
SQ SEQUENCE 51 AA; 5836 MW; 44E0556203675888 CRC64;

Query Match 13.2%; Score 35; DB 1; Length 51;
Best Local Similarity 26.1%; Pred. NO. 6.7e+02;
Matches 12; Conservative 6; Mismatches 22; Indels 6; Gaps 1;

QY 5 YGLRPGSSGSPS-----LDEKKIAKMEKASVFNVNSSGSPSLHW 44
Db 3 YGVYRPSGVECPGILQAQOIEYLAINGVYGAKLVESEDLHW 48

RESULT 9

RS19_YEREN
ID RS19_YEREN STANDARD; PRT; 32 AA.
AC Q56847;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 30S ribosomal protein S19 (Fragment).
GN RPS5.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=6471/76 / SROTYPE O:3;
RX MEDLINE=96382736; PubMed=8790600;
RA Mertz A.K.H., Daser A., Skurnik M., Wiesmuller K., Braun J., Appel H.,
RA Batsford S., Wu P., Distler A., Sieper J.;
RT "The evolutionarily conserved ribosomal protein L23 and the cationic
RT urease beta-subunit of Yersinia enterocolitica O:3 belong to the
RT immunodominant antigens in Yersinia-triggered reactive arthritis:
RT implications for autoimmunity.";
RL Mol. Med. 1:44-55(1994).
CC -----

CC -!- FUNCTION: PROTEIN S19 FORMS A COMPLEX WITH S13 THAT BINDS STRONGLY

CC TO THE 16S RIBOSOMAL RNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U11251; AAC43514.1; -
DR InterPro; IPR002222; Ribosomal_S19.
DR Pfam; PF00203; Ribosomal_S19; 1.
DR ProDom; PD001012; Ribosomal_S19; 1.
DR PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW Ribosomal protein; rRNA-binding.
FT NON_TER 32 32
SQ SEQUENCE 32 AA; 3659 MW; E534F701330F0338 CRC64;

Query Match 12.8%; Score 34; DB 1; Length 32;
Best Local Similarity 46.7%; Pred. NO. 5.3e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 13 GPSLDEKKIAKMEKA 27
Db 8 GPFIDLHLKKVEKA 22

RESULT 10

Y16N_BP74
ID Y16N_BP74 STANDARD; PRT; 42 AA.
AC P39249; Q9T034;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 5.0 kDa protein in ndd-denB intergenic region.
GN Y16N OR NDD.4.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT "Bacteriophage T4 genome analysis.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF158101; AAD42620.1; -
KW Hypothetical protein.
SQ SEQUENCE 42 AA; 4954 MW; C49805F365992E59 CRC64;

Query Match 12.8%; Score 34; DB 1; Length 42;
Best Local Similarity 42.1%; Pred. NO. 7.2e+02;
Matches 8; Conservative 7; Mismatches 2; Indels 2; Gaps 1;

QY 16 LDEKKIAKMEKASVFNV 34
Db 1 MEEKMKKIIKA--IWNV 17

RESULT 11

PRVM_SCYCA
ID PRVM_SCYCA STANDARD; PRT; 50 AA.


```
AC P35491.1
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE Parvalbumin (Fragment).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RC TISSUE=Spinal cord;
RX MEDLINE=92319231; PubMed=1620290;
RA Conlon J.M., O'Harte F., Smith D.D., Balmert R.J., Hazon N.;
RT "Purification and characterization of urotensin II and parvalbumin
from an elasmobranch fish, Scyliorhinus canicula (common dogfish).";
RL Neuroendocrinology 55:230-235(1992).
CC -!- FUNCTION: PROBABLY REGULATES THE ACTIVITY OF THE CAUDAL
CC NEUROSECRETORY SYSTEM.
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS TWO CALCIUM IONS (BY
CC SIMILARITY).
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE PARVALBUMINS SUBFAMILY.
DR HSSP: P30563; SPAL.
DR InterPro: IPR002048; EF-hand.
DR PROSITE: PS00018; EF_HAND; PARTIAL.
KW Calcium-binding; Muscle protein; Repeat.
FT NON_TER 50
SQ SEQUENCE 50 AA; 5558 MW; 6E4A2D264A312DBE CRC64;

Query Match 12.8%; Score 34; DB 1; Length 50;
Best Local Similarity 29.4%; Pred. No. 8.7e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 12; Gaps 2;

QY 9 PGSSGSPSLDEKKIAKM-----EKASSVFNVV 34
DB 21 FG-----SFDHKFFQLVGLKGKTHEQVKVFNIL 50

RESULT 12
ID PP71_HCMVT STANDARD; PRT; 28 AA.
AC P24429;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE 71 kDa upper matrix phosphoprotein (pp71) (Fragment).
GN UL82.
OS Human cytomegalovirus (strain Towne).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10363;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91220654; PubMed=1850902;
RA Pande H., Campo K., Tanamachi B., Zaia J.A.;
RT "Human cytomegalovirus strain Towne pp65 gene: nucleotide sequence
and expression in Escherichia coli.";
RL Virology 182:220-228(1991).
CC -!- SIMILARITY: BELONGS TO THE UL82 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M67443; AAA45995.1; -.
KW Phosphorylation; Matrix protein.
FT NON_TER 28
```

```
SQ SEQUENCE 28 AA; 2554 MW; 3C96A69C384425BD CRC64;

Query Match 12.6%; Score 33.5; DB 1; Length 28;
Best Local Similarity 45.0%; Pred. No. 5.3e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 9 PGSSGSPSLDEKKIAKMEKAS 28
DB 8 PG-EGPSSAAALISEAAAS 26

RESULT 13
Y799_TREPA
ID Y799_TREPA STANDARD; PRT; 50 AA.
AC O8377;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0799.
GN TP0799.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001250; AAC65768.1; -.
DR TIGR; TP0799; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 50 AA; 5796 MW; 6E788CB5CC9A818F CRC64;

Query Match 12.6%; Score 33.5; DB 1; Length 50;
Best Local Similarity 37.5%; Pred. No. 1e+03;
Matches 9; Conservative 3; Mismatches 5; Indels 7; Gaps 1;

QY 23 KMEKASSVFNVNSSGSPSLHWSY 46
DB 29 KEENATGLMN-----PSLHTSH 45

RESULT 14
NPF_ARTTR
ID NPF_ARTTR STANDARD; PRT; 36 AA.
AC P41334;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuropeptide F (NPF).
OS Artiposthia triangulata.
OC Eukaryota; Metazoa; Platyhelminthes; Turbellarian platyhelminths;
OC Rhabditophora; Seriata; Tricladida; Terricola; Geoplanidae;
OC Arthurdendyus.
OX NCBI_TaxID=132421;
```

```
RN [1]
RX MEDLINE=92362227; PubMed=1354101;
RA Curry W.J., Shaw C., Johnston C.F., Thim L., Buchanan K.D.;
RT "Neuropeptide F: primary structure from the tubellarian, Artioposthia
RL triangulata.";
RL Comp. Biochem. Physiol. 101C:269-274(1992).
CC -1- FUNCTION: MAY PERFORM AN IMPORTANT NEUROTRANSMITTER FUNCTION AND
CC MAY REGULATE MUSCULAR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: CENTRAL AND PERIPHERAL NERVOUS SYSTEM, AND
CC MUSCULAR PHARYNX.
CC -1- SIMILARITY: BELONGS TO THE NPY / PPY / PYY FAMILY.
DR InterPro; IPR001955; Pancreat_hormn.
DR Pfam; PF00159; hormone3; 1.
DR SMART; SM00309; PAH; 1.
DR PROSITE; PS00265; PANCREATIC_HORMONE_1; 1.
DR PROSITE; PS00276; PANCREATIC_HORMONE_2; 1.
KW Neuropeptide; Amidation.
FT MOD_RES 36 36 AMIDATION.
SQ SEQUENCE 36 AA; 4433 MW; D959E4C464B7E70C CRC64;

Query Match 12.5%; Score 33; DB 1; Length 36;
Best Local Similarity 60.0%; Pred. No. 8.1e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 LRGSGGSPSLDEKKI 21
   ||| ||| ||| |||
DB 5 LRPRSFSEDEYQI 19

RESULT 15
TPX2_STRPN
ID TPX2_STRPN STANDARD; PRT; 49 AA.
AC R42365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable thiol peroxidase (EC 1.11.1.-) (Fragment).
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R36A;
RX MEDLINE=94086122; PubMed=7505262;
RA Sampson J.S., O'Connor S.P., Stinson A.R., Tharpe J.A., Russell H.;
RT "Cloning and nucleotide sequence analysis of psaA, the Streptococcus
RT pneumoniae gene encoding a 37-kilodalton protein homologous to
RT previously reported Streptococcus sp. adhesins.";
RL Infect. Immun. 62:319-324(1994).
CC -1- FUNCTION: HAS ANTIOXIDANT ACTIVITY. COULD REMOVE PEROXIDES OR
CC H(2)O(2) (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY. TPX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; L19055; AAA16799.1; -
DR InterPro; IPR02065; TPX.
DR PROSITE; PS01265; TPX; PARTIAL.
KW Oxidoreductase; Peroxidase.
FT INIT_MET 0 0 BY SIMILARITY.
FT NON_TER 49 49
SQ SEQUENCE 49 AA; 5224 MW; 4F81A63DD548B33 CRC64;

Query Match , 12.5%; Score 33; DB 1; Length 49;
```

```
Best Local Similarity 28.6%; Pred. No. 1.1e+03;
Matches 6; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 14 PSLDEKKIAKMEKASSVFNVV 34
   |::|::| |::|
DB 29 PNLEKSLADFAFGKKKVLVI 49

Search completed: October 10, 2002, 23:11:13
Job time : 14 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 23:07:11 ; Search time 29 Seconds
(without alignments)

304.232 Million cell updates/sec

Title: US-09-848-834a-20

Perfect score: 265

Sequence: 1 XHSWYGLRPGSGPSLDEKK.....NVVSSSGPSLHWSYGLRPX 51

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 30812

Minimum DB seq length: 0

Maximum DB seq length: 51

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	17.7	50	5 Q9N416	Q9N416 caenorhabdi
2	45	17.0	33	13 Q9PT34	Q9PT34 oncorhynch
3	45	17.0	33	13 Q9W7G0	Q9W7G0 oncorhynch
4	42.5	16.0	48	11 Q923S0	Q923S0 mus musculu
5	40	15.1	30	16 Q50832	Q50832 borrelia bu
6	39.5	14.9	30	2 Q9R5S3	Q9R5S3 leuconostoc
7	38	14.3	20	11 Q9EQX8	Q9EQX8 mus musculu
8	38	14.3	30	4 Q9UM16	Q9UM16 homo sapien
9	38	14.3	32	5 Q9U340	Q9U340 caenorhabdi
10	38	14.3	44	8 Q32439	Q32439 hordeum vul
11	38	14.3	47	6 Q29327	Q29327 sus scrofa
12	37.5	14.2	42	10 Q9S8U3	Q9S8U3 chlamydomon
13	37.5	14.2	50	16 Q916X1	Q916X1 pseudomonas
14	37	14.0	27	4 Q16162	Q16162 homo sapien
15	37	14.0	39	4 Q96CW4	Q96CW4 homo sapien
16	37	14.0	50	5 Q95ZB1	Q95ZB1 leishmania

ALIGNMENTS

RESULT 1

Q9N416 ID Q9N416 PRELIMINARY; PRT; 50 AA.
AC Q9N416; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 5.3 KDA PROTEIN.
GN Y39A3CL.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Kalicki J., Smith A.;
RT "The sequence of C. elegans cosmid Y39A3CL.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024763; AAF60518.1;
KW Hypothetical protein.
SQ SEQUENCE 50 AA; 5258 MW; 4BA34E5DA85FF21F CRC64;
Query Match 17.7%; Score 47; DB 5; Length 50;
Best Local Similarity 39.1%; Pred. NO. 62;

Q9aeu8 brassica na
Q9aeu6 brassica ol
Q9we5 rattus sp.
Q9azh4 bacterioph
Q9gk18 sus scrofa
O19186 ovis aries
Q9tv94 bos taurus
Q9n278 pan troglod
O86256 klebsiella
Q9kvx1 vibrio chol
P90686 amphipholis
P95809 streptococc
Q9nm75 leishmania
Q9riu5 staphylococ
Q41781 zea mays (m
Q9k201 chlamydia p
O82954 burkholderi
O27973 archaeoglob
Q9bw55 homo sapien
Q9vj93 dirosophila
Q9ezk7 ehrlichia e
Q967x1 folsonia ca
Q91k37 hepatitis c
O42780 ustilago ho
Q9t0s4 bacterioph
P95807 streptococc
Q9kpf7 vibrio chol
Q9ka16 bacillus ha
Q9zg37 chlamydia t

RA	Von Schalburg K.R., Sherwood N.M.;
RT	"Regulation and expression of gonadotropin-releasing hormone gene
RN	differs in brain and gonads in rainbow trout.";
RL	Endocrinology 140:3012-3024(1999).
[2]	
RP	SEQUENCE FROM N.A.
von Schalburg K.R., Sherwood N.M.;	
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.	
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY	
SIMILARITY).	
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.	
EMBL: AF110993; AAD43463.1; -.	
InterPro: IPR002012; GnRH.	
Pfam: PF00446; GnRH; 1.	
DR PROSITE; PS00473; GNRH; 1.	
DR Amidation; Hormone.	
FT NON_TER 33	
SQ SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;	
Query Match 17.0%; Score 45; DB 13; Length 33;	
Best Local Similarity 77.8%; Pred. No. 70;	
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY 2 HWSYGRLRPG 10	
DB 25 HWSYGLWLPG 33	
RESULT 4	
Q923S0 PRELIMINARY; PRT; 48 AA.	
ID Q923S0	
AC Q923S0;	
DT 01-DEC-2001 (TrEMBLrel. 19, Created)	
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE STEROIDIN 1 (FRAGMENT).	
GN ABCG5.	
OS Mus musculus (Mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX NCBI_TaxID=10090;	
RN [1]	
RP SEQUENCE FROM N.A.	
Lu K., Lee M., Patel S.B.;	
RC STRAIN=C57BL/6;	
RT "Molecular cloning, genomic structure and characterization of novel	
murine ABC genes Abcg5 and Abcg8."	
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.	
EMBL: AF404108; AAR85590.1; -.	
FT NON_TER 48	
SQ SEQUENCE 48 AA; 5019 MW; C2C512DBF8A82D20 CRC64;	
Query Match 16.0%; Score 42.5; DB 11; Length 48;	
Best Local Similarity 28.2%; Pred. No. 2,3e+02;	
Matches 11; Conservative 8; Mismatches 19; Indels 1; Gaps 1;	
QY 10 GSSGPSLEKKIAKWEKASSVFNVVNSSSGPSLHWSYGL 48	
: : : : :	
DB 11 GARGPHINRGSLSEQGSVTGTGARHSLG-VLHVSVSV 48	
RESULT 5	
I050832 PRELIMINARY; PRT; 30 AA.	
ID O50832	
AC O50832;	
DT 01-JUN-1998 (TrEMBLrel. 06, Created)	
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)	
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)	
DE HYPOTHETICAL 3.6 KDA PROTEIN.	
GN BIK29.	
OS Borrelia burgdorferi (Lyme disease spirochete).	
OG plasmid lp36.	
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.	

```

OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J.R., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA Van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL: AF000788; AAC66163.1; -.
KW TIGR; BBK29; -.
KW Hypothetical protein; Plasmid: Complete proteome.
SQ SEQUENCE 30 AA; 3575 MW; B36FLC321118A2B6 CRC64;

Query Match 15.1%; Score 40; DB 16; Length 30;
Best Local Similarity 33.3%; Pred. No. 2.8e+02;
Matches 8; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 10 GSSGSPSLDEKKIAKMEKASSVFNV 33
Db 2 GNKPSSGDKKMKIRKLSSYYKI 25

RESULT 6
QYR5S3 ID Q9R5S3 PRELIMINARY; PRT; 30 AA.
AC Q9R5S3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE D-LACTATE DEHYDROGENASE.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE.
RX MEDLINE=92031480; PubMed=1931965;
RA Bugg T.D., Wright G.D., Dutka-Malen S., Arthur M., Courvalin P.,
RA Walsh C.T.;
RT "Molecular basis for vancomycin resistance in Enterococcus faecium BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA.";
RL Biochemistry 30:10408-10415(1991).
DR HSSP; P26297; IDLD.
SQ SEQUENCE 30 AA; 3544 MW; B1696A5C11B2038C CRC64;

Query Match 14.9%; Score 39.5; DB 2; Length 30;
Best Local Similarity 43.5%; Pred. No. 3.3e+02;
Matches 10; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 3 WSYGLRPGSGSLDEKKIAKME 25
Db 4 FAYGIRDDER-PSLEENKAANPE 25

RESULT 7
QYEQX8 ID Q9EQX8 PRELIMINARY; PRT; 20 AA.
AC Q9EQX8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GLUTAMATE DECARBOXYLASE 65 (FRAGMENT).
GN GAD65.
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Makinae K., Kobayashi T., Kobayashi T., Sinkawa H., Sakagami H.,
RA Kondo H., Tashiro F., Miyazaki J., Obata K., Tamura S., Yanagawa Y.;
RT "Structure of the Mouse Glutamate Decarboxylase 65 Gene and its Promoter.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032757; BAB20415.1; -.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 1989 MW; 28F544B0E4C3768D CRC64;

Query Match 14.3%; Score 38; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 WSYGLRPGSSGP 14
Db 9 WSGSGEDGSADP 20

RESULT 8
QYUMI6 ID Q9UMI6 PRELIMINARY; PRT; 30 AA.
AC Q9UMI6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OCT-2 FACTOR (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89107992; PubMed=3265124;
RA Clerc R.G., Corcoran L.M., Lebowitz J.H., Baltimore D., Sharp P.A.;
RT "The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains.";
RL Genes Dev. 2:1570-1581(1988).
DR EMBL: M36772; AAA36392.1; -.
FT NON_TER 1 1
SQ SEQUENCE 30 AA; 3087 MW; 8368E8D240E28762 CRC64;

Query Match 14.3%; Score 38; DB 4; Length 30;
Best Local Similarity 50.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 34 VNSSSGPSLHWS 45
Db 13 LNPSTGPGLWN 24

RESULT 9
QYU340 ID Q9U340 PRELIMINARY; PRT; 32 AA.
AC Q9U340;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE W06G6.9 PROTEIN.
GN W06G6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
SQ SEQUENCE 32 AA; 3087 MW; 8368E8D240E28762 CRC64;

```

RN SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z83129; CAB63325.1; -.
SQ SEQUENCE 32 AA; 3775 MW; E49D031288C6806A CRC64;

Query Match 14.3%; Score 38; DB 5; Length 32;
Best Local Similarity 41.7%; Pred. No. 5.6e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 HWSYGLRPGSSG 13
| | : | | | |
Db 18 HWQNAKPGEWG 29

RESULT 10
Q32439 PRELIMINARY; PRT; 44 AA.
AC Q32439;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF44.
OS Hordeum vulgare (Barley).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90291518; PubMed=1694111;
RA Sexton T.B., Jones J.T., Mullet J.E.;
RT "Sequence and transcriptional analysis of the barley cDNA region
RT upstream of psbD-psbC encoding trnK(UUU), rps16, trnQ(UUG), psbK,
RT psbI, and trnS(GCU)."
RL Curr. Genet. 17:445-454(1990).
DR EMBL; X52765; CAA36977.1; -.
KW Chloroplast.
SQ SEQUENCE 44 AA; 5188 MW; 7A11B128EB8DAE4C CRC64;

Query Match 14.3%; Score 38; DB 8; Length 44;
Best Local Similarity 43.5%; Pred. No. 8.1e+02;
Matches 10; Conservative 1; Mismatches 8; Indels 4; Gaps 1;
QY 3 WSYGLRPGSSGSLDEKKIAKME 25
| | | | | | | | | |
Db 19 WRPGFRPG-----SLDKNPKIKSE 37

RESULT 11
Q29327 PRELIMINARY; PRT; 47 AA.
AC Q29327;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF PROTEIN (FRAGMENT).
GN ORF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA

library."
RL Mamm. Genome 7:509-517(1996).
DR EMBL; F15092; CAA23348.1; -.
FT NON_TER 1
FT NON_TER 47
SQ SEQUENCE 47 AA; 5334 MW; DF1532AD980817E9 CRC64;
Query Match 14.3%; Score 38; DB 6; Length 47;
Best Local Similarity 28.6%; Pred. No. 8.8e+02;
Matches 8; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
QY 3 WSYGLRPGSSGSLDEKKIAKMEKASSV 30
| | : | | | | | | | |
Db 11 WCHSLOPGRTPWISGRKRTMTXTSMV 38
RESULT 12
Q9S8U3 PRELIMINARY; PRT; 42 AA.
AC Q9S8U3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PHOTOSYSTEM I LIGHT-HARVESTING COMPLEX CHLOROPHYLL A/B PROTEIN, P14.1
DE (FRAGMENT).
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE.
RX MEDLINE=93100280; PubMed=1464588;
RA Bassi R., Soen S.Y., Frank G., Zuber H., Rochaix J.D.;
RT "Characterization of chlorophyll a/b proteins of photosystem I from
RT Chlamydomonas reinhardtii."
RL J. Biol. Chem. 267:25714-25721(1992).
DR InterPro; IPR001344; Chloro_AB_bind.
DR Pfam; PF00504; chloroa_b_bind; 1.
SQ SEQUENCE 42 AA; 4383 MW; AGAC4D3B2FF9A5A8 CRC64;
Query Match 14.2%; Score 37.5; DB 10; Length 42;
Best Local Similarity 30.8%; Pred. No. 8.9e+02;
Matches 12; Conservative 6; Mismatches 16; Indels 5; Gaps 1;
QY 17 DEKKIAKMEKASSVFNVNSSSG-----PSLHWSYGLRP 50
| | | | | | | | | | | | | | | | | | | | | |
Db 1 EEKSIKVDKSLQDLAVGASQSSSLAYLDGSLPGDFGDP 39
RESULT 13
Q9I6X1 PRELIMINARY; PRT; 50 AA.
AC Q9I6X1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN PA0161.
GN PA0161.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 23:05:56 ; Search time 27 Seconds

(without alignments)
209.806 Million cell updates/sec

Title: US-09-848-834A-20

Perfect score: 265

Sequence: 1 XHSYGLRPGSGPSLDEKK.....NVNSSSGPSLHWSYGLRPX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 355628

Minimum DB seq length: 0

Maximum DB seq length: 51

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127.5	48.1	33	AA62715	LHRH-containing im
2	97.5	36.8	49	AAW03944	GnRH 4-repeat sequ
3	97.5	36.8	49	AAW79567	GnRH-2. Synthetic
4	97.5	36.8	49	AAW61542	Peptide hormone Gn
5	97.5	36.8	49	AAV58363	Four-copy gonadotr
6	97.5	36.8	49	AAV58135	GnRH analogue mult
7	90.5	34.2	42	AAV20865	GnRH tandem dimer
8	90	34.0	23	AAW20864	GnRH tandem repeat
9	89.5	33.8	33	AAW63663	Peptide comprising
10	89.5	33.8	33	AAW63516	A peptide which ma
11	89	33.6	19	AAW98951	Vaccine related MH

12	89	33.6	21	10	AA62715	Sequence of modifi
13	89	33.6	21	16	AA62715	Plasmodium falcipa
14	89	33.6	21	16	AA62715	Malaria circumspor
15	89	33.6	21	16	AA62715	P. falciparum CS p
16	89	33.6	21	16	AA62715	Malaria circumspor
17	89	33.6	21	17	AAW05612	Circumsporozoite h
18	89	33.6	21	18	AAW35440	T-cell stimulatory
19	89	33.6	21	20	AAV23252	Peptide derived fr
20	89	33.6	21	21	AAV80071	Pathogen derived f
21	89	33.6	21	21	AAV54553	T helper cell (Th)
22	89	33.6	21	21	AAV58777	Unidentified pepti
23	89	33.6	21	22	AAV99706	Plasmodium falcipa
24	89	33.6	21	22	AAV99706	Plasmodium falcipa
25	89	33.6	21	22	AAV99706	Plasmodium falcipa
26	89	33.6	21	22	AAV99706	Plasmodium falcipa
27	89	33.6	21	22	AAV99706	Plasmodium falcipa
28	89	33.6	21	22	AAV99706	Plasmodium falcipa
29	89	33.6	21	22	AAV99706	Plasmodium falcipa
30	89	33.6	21	22	AAV99706	Plasmodium falcipa
31	89	33.6	21	22	AAV99706	Plasmodium falcipa
32	88.5	33.4	30	11	AAV07323	Luteinising hormon
33	88.5	33.4	40	20	AAV31183	Ubiquitin fusion p
34	88.5	33.4	40	22	AAV31183	GnRH sequence #2.
35	88.5	33.4	41	20	AAV31182	Ubiquitin fusion p
36	88.5	33.4	41	22	AAV31182	GnRH sequence #1.
37	87.5	33.0	40	21	AAV96085	Cattle gonadotropi
38	85	32.1	18	21	AAV49259	CD4+ T cell epitop
39	84.5	31.9	20	19	AAW47438	Antigenic peptide.
40	84.5	31.9	20	20	AAV31174	Ubiquitin fusion p
41	84.5	31.9	20	20	AAV31177	Ubiquitin fusion p
42	84.5	31.9	20	20	AAV31178	Ubiquitin fusion p
43	84.5	31.9	20	20	AAV31179	Ubiquitin fusion p
44	84.5	31.9	20	22	AAV71945	GnRH dimer. Unide
45	84.5	31.9	20	22	AAV71948	GnRH dimer peptide

ALIGNMENTS

RESULT 1
AA62715
ID AA62715 standard; peptide; 33 AA.
XX AA62715;
AC AA62715;
XX
DT 10-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
KW Helper T cell epitope; universal immune stimulator; Invasin; haptens;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility;
KW Plasmodium falciparum circumsporozoite.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1..21 /note= "Plasmodium falciparum circumsporozoite
FT Domain 24..33 helper T cell epitope"
FT Domain /note= "LHRH haptens"
XX
PN WO9425060-A.
PD 10-NOV-1994.
XX
PF 28-APR-1994; 94WO-US04832.
XX
PR 27-APR-1993; 93US-0057166.
PR 14-APR-1994; 94US-0229275.
XX
PA (LADD/) LADD A E.

PA (WANG/) WANG C Y.
 XX (ZAMB/) ZAMB T.
 PI Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX
 XX Claim 8; Page 86; 213pp; English.
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX
 XX Sequence 33 AA;
 SQ
 Query Match 48.1%; Score 127.5; DB 15; Length 33;
 Best Local Similarity 81.8%; Pred. No. 1.3e-09;
 Matches 27; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
 QY 18 EKKIARMEKASSVFNVSNGSPSLHWSYGLRP 50
 DB 3 EKKIARMEKASSVFNVSNGS---HWSYGLRP 32
 RESULT 2
 AAW03944
 ID AAW03944 standard; Protein; 49 AA.
 AC AAW03944;
 XX
 XX 20-NOV-1996 (first entry)
 DT
 XX GnRH 4-repeat sequence.
 DE
 XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
 KW fusion protein; immunogen; vaccine; fertility control;
 KW contraceptive; sterilisation; PCB113; PCB11.
 XX
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..10
 FT /label= GnRH
 FT Peptide 11..13
 FT /label= Spacer
 FT Peptide 14..23
 FT /label= GnRH
 FT Peptide 24..26
 FT /label= Spacer
 FT Peptide 27..36
 FT /label= GnRH
 FT Peptide 37..39
 FT /label= Spacer
 FT Peptide 40..49

FT /label= GnRH
 XX WO9624675-A1.
 PN 15-AUG-1996.
 XX
 PD 24-JAN-1996; 96WO-CA00049.
 XX
 PF 10-FEB-1995; 95US-0387156.
 XX
 PR (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Manns JG, Potter AA;
 XX WPI; 1996-384447/38.
 DR N-PSDB; AAT37178.
 XX
 XX Gonadotropin-releasing hormone multimer fusion proteins - with
 PT leukotoxin polypeptide for increased immunogenicity, useful in
 PT antifertility vaccine prodn.
 XX
 XX Example 2; Fig 1B; 87pp; English.
 PS
 XX A synthetic DNA sequence (AAT37178) codes for a gonadotropin
 CC releasing hormone (GnRH) tetramer (AAW03944), in which the 4
 CC repeat units are separated by spacers designed to increase
 CC immunogenicity. The DNA sequence was incorporated into vector
 CC pAA352 (ATCC 68283), which contains a truncated leukotoxin
 CC gene (LKT 352) derived from Pasteurella haemolytica, to
 CC give plasmid pCB113 (AAT37176). Escherichia coli transformants
 CC produce an LKT-GnRH fusion protein (see also AAW03942 and AAW03943)
 CC useful as a vaccine for fertility control of domestic or farm animals.
 XX
 XX Sequence 49 AA;
 SQ
 Query Match 36.8%; Score 97.5; DB 17; Length 49;
 Best Local Similarity 42.9%; Pred. No. 1.7e-05;
 Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;
 QY 2 HWSYGLRPGSGSLDEKIAKMEKASSVFNVSNGSPSLHWSYGLRP 50
 DB 2 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 35
 RESULT 3
 AAW79567
 ID AAW79567 standard; Protein; 49 AA.
 XX
 XX AAW79567;
 AC
 XX 24-DEC-1998 (first entry)
 DT
 XX GnRH-2.
 DE
 XX Gonadotropin releasing hormone; GnRH; decapeptide; hypothalamus; spacer;
 KW LH; luteinising hormone; FSH; follicle stimulating hormone; vertebrate;
 KW pyroglu; chimera; leukotoxin polypeptide; multimer; vaccine; tumour;
 KW immunogenic.
 XX
 XX Synthetic.
 OS
 XX WO9806848-A1.
 PN 19-FEB-1998.
 PD
 XX 08-AUG-1997; 97WO-CA00559.
 PF
 XX 09-AUG-1996; 96US-0694865.
 PR
 XX (UYSA-) UNIV SASKATCHEWAN.
 PA
 PI Manns JG, Potter AA;
 XX

CC in vivo; or used in therapy of certain autoimmune diseases. The
CC pseudopeptides are useful as carriers for antigens or other therapeutic
CC agents due to their ability to form non-covalent bonds via the
CC hydrophobic or hydrophilic auxiliary spacer. The present sequence
CC represents a peptide, which has conjugation sites for the pseudopeptides
CC of the invention.

XX SQ Sequence 33 AA;

Query Match 33.8%; Score 89.5; DB 22; Length 33;
Best Local Similarity 76.9%; Pred. No. 0.00012;
Matches 20; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 11 SSGPSLDEKKIAKMEKASSVFNVNS 36
:: : |||||
Db 9 NANPDI-EKKIAKMEKASSVFNVNS 33

RESULT 10
AAG63516
ID AAG63516 standard; peptide; 33 AA.

XX AC AAG63516;

XX DT 15-OCT-2001 (first entry)

XX DE A peptide which may be conjugated to pseudopeptides.

XX KW Pseudopeptide; immunomodulation; adjuvant; antigen presenting cell;
XX KW macrophage; dendritic cell; cytokine production; immunocompetent cell;
XX KW autoimmune disease.

XX OS Synthetic.

XX PN WO200146126-A1.

XX PD 28-JUN-2001.

XX PF 21-DEC-2000; 2000WO-FR03650.

XX PR 22-DEC-1999; 99WO-IB02038.

XX PA (OMPH-) OM-PHARMA.

XX PI Bauer J, Martin OR, Rodriguez S;

XX DR WPI; 2001-496651/54.

XX PT New amphiphilic acylated pseudopeptides having a functionalized
XX PT auxiliary spacer, useful as immunomodulators, e.g. as adjuvants in
XX PT vaccines

XX PS Example 3.4; Page 89; 267pp; French.

XX CC The specification describes N-Acylated pseudopeptides, which have
XX CC a neutral or charged acidic group at one terminal and a functionalized
XX CC auxiliary spacer at the other. The pseudopeptides show immunomodulatory
XX CC and adjuvant action, based on activation of antigen presenting cells
XX CC (e.g. macrophages or dendritic cells), induction of differentiation of
XX CC dendritic cells, induction of cytokine production and induction of
XX CC maturation of immunocompetent cell strains originating from hematopoietic
XX CC and lymphoid organs. They reinforce humoral and cellular immunity. They
XX CC can be grafted onto antigens (to modulate immune response) or onto
XX CC drugs (to improve the therapeutic activity or targeting). The
XX CC pseudopeptides are thus useful in human or veterinary medicine as
XX CC immunizing or diagnostic agents. Typically, the pseudopeptides are used
XX CC as adjuvants together with (or covalently bonded to) antigens for
XX CC vaccination against viral, parasitic/protozoal, microbial or fungal
XX CC infections; incubated with blood cells ex vivo, to render the cells
XX CC immunocompetent before reintroduction in vivo; or used in therapy of
XX CC certain autoimmune diseases. The present sequence represents a
XX CC peptide which may be conjugated to pseudopeptides of the invention.

SQ Sequence 33 AA;
Query Match 33.8%; Score 89.5; DB 22; Length 33;
Best Local Similarity 76.9%; Pred. No. 0.00012;
Matches 20; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 11 SSGPSLDEKKIAKMEKASSVFNVNS 36
:: : |||||
Db 9 NANPDI-EKKIAKMEKASSVFNVNS 33

RESULT 11
AAM98951
ID AAM98951 standard; Peptide; 19 AA.

XX AC AAM98951;

XX DT 07-DEC-2001 (first entry)

XX DE Vaccine related MHC ligand peptide SEQ ID NO:54.

XX KW Glutamic acid; glutamine; vaccine; major histocompatibility complex;
XX KW MHC; immunomodulator; antiallergic; endocrine; neuroprotectant;
XX KW virucidal; bactericidal; antiparasitic; fungicidal; cytostatic;
XX KW medicine; pharmaceutical; immune disorder; immune deficiency;
XX KW autoimmune; hypersensitivity; allergy; graft rejection; infection;
XX KW hormonal disorder; central nervous system disease; cancer; melanoma;
XX KW anti-melanoma vaccine; human immunodeficiency virus.

XX OS Plasmodium malariae.

XX PN WO200170772-A2.

XX PD 27-SEP-2001.

XX PF 22-MAR-2001; 2001WO-FR00872.

XX PR 23-MAR-2000; 2000FR-0003711.

XX PA (FABR) FABRE MEDICAMENT SA PIERRE.

XX PI Klinguer-Hamour C, Corvaia N, Beck A, Goetsch L;

XX DR WPI; 2001-611470/70.

XX PT Stabilized pharmaceutical containing N-terminal glutamic acid or
XX PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
XX PT with strong acid

XX PS Claim 9; Page 39; 149pp; French.

XX CC The present invention describes a pharmaceutical compound (I) that
XX CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue
XX CC in the form of an addition salt with a strong, physiologically
XX CC acceptable acid (II). Also described are: (a) a pharmaceutical
XX CC composition containing at least one (I); (b) a vaccine containing at
XX CC least one (I) where this is a major histocompatibility complex (MHC)
XX CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
XX CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
XX CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
XX CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
XX CC fungicidal and cytostatic activities. (I) are useful, in human or
XX CC veterinary medicine, in pharmaceutical compositions (for treating immune
XX CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
XX CC allergy, graft rejection, infection, hormonal disorders and central
XX CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
XX CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
XX CC or fungal infections; or (ii) of cancers. A particular application is in
XX CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
XX CC diseases associated with interactions between MHC and (I), e.g. melanoma
XX CC and human immunodeficiency virus infection. AAM98951 to AAM99592
XX CC represent peptides which can be used in pharmaceutical compounds from
XX CC the present invention.

```

XX SQ Sequence 19 AA;
Query Match 33.6%; Score 89; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAKMEKASSVFNVNS 36
    | | | | | | | | | | | | | | | | | | | |
Db 1 EKKIAKMEKASSVFNVNS 19

RESULT 12
AAP91504
ID AAP91504 standard; peptide; 21 AA.
XX AC AAP91504;
XX DT 13-MAR-1992 (first entry)
XX PF Sequence of modified Plasmodium circumsporozoite T-cell epitope CS.T3.
XX DE Circumsporozoite peptide; T-cell epitope; immunogenic composition;
XX KW vaccine.
XX OS Plasmodium falciparum.
XX FH Key Location/Qualifiers
FT Misc-difference 1..2 /note= "May be H-Asp-Ile, H-Ile, or H-"
FT FT Misc-difference 19..21 /note= "May be Val-Asn-Ser-OH, Val-Asn-OH,
FT Val-OH or -OH "
XX PN EP343460-A.
XX PD 29-NOV-1989.
XX PF 12-MAY-1989; 89EP-0108618.
XX PR 24-MAY-1988; 88GB-0012214.
XX PA (HOFF ) HOFFMANN-LA ROCHE AG.
XX PI Sinigaglia F;
XX DR WPI; 1989-349561/48.
XX PT Modified Plasmodium CS peptide - used as a universally recognised
PT T-cell epitope in vaccines to elicit an immune response against
PT pathogenic agents
XX PS Claim 1; Page 16; 23pp; English.
XX CC Peptide CS.T3 (residues 1-21) corresponds to residues 378-398 of the
CC CS protein from P. falciparum but contains 2 Ala residues in place
CC of the native protein's Cys residues at positions 384 and 389. Also
CC claimed is AAP91504 (or modified forms, see FT) associated with an
CC antigenic structure representing a B-cell epitope, pref. a multiple
CC antigenic peptide, esp. multimers of the repeat sequences NANP
CC present in P. falciparum CS protein.
XX SQ Sequence 21 AA;
Query Match 33.6%; Score 89; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAKMEKASSVFNVNS 36
    | | | | | | | | | | | | | | | | | | | |
Db 3 EKKIAKMEKASSVFNVNS 21

XX SQ Sequence 21 AA;
Query Match 33.6%; Score 89; DB 16; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAKMEKASSVFNVNS 36
    | | | | | | | | | | | | | | | | | | | |
Db 3 EKKIAKMEKASSVFNVNS 21

```

```

RESULT 13
AAR82586
ID AAR82586 standard; peptide; 21 AA.
XX AC AAR82586;
XX DT 13-JUN-1996 (first entry)
XX PF Plasmodium falciparum circumsporozoite helper T cell epitope, PF.
XX DE IGE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
XX KW vaccine; allergy; antibody; constant heavy chain.
XX OS Plasmodium falciparum.
XX PN WO9526365-A1.
XX PD 05-OCT-1995.
XX PF 24-MAR-1995; 95WO-US03741.
XX PR 25-OCT-1994; 94US-0328912.
XX PR 28-MAR-1994; 94US-0218461.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PT Wang CY;
XX DR WPI; 1995-351297/45.
XX PT Synthetic peptide-based immunogen contg. IGE CH4 peptide and helper
FT T cell epitope - useful for eliciting antibody prodn. for allergy
FT treatment
XX PS Claim 3; Page 23; 87pp; English.
XX CC AAR82571-91 are helper T cell epitopes which can be used in the
CC preparation of a peptide immunogen that is useful in vaccines for
CC treating allergic reactions. In the immunogen an IGE CH4 peptide
CC is attached C-terminally to a series of amino acids including a
CC helper T cell epitope. The immunogen may also opt. contain a fatty
CC acid or fatty acid derivative, an invasin domain or alpha-NH2. The
CC immunogen produces high titres of antibodies to the effector site
CC in human IGE heavy chain (the CH4 domain peptide) which inhibit mast
CC cell activation and reduce allergen-induced IGE prodn. The immunogens
CC may be used in either a radially branching multimeric form or a
CC linearly arranged monomeric form.
XX SQ Sequence 21 AA;
Query Match 33.6%; Score 89; DB 16; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 EKKIAKMEKASSVFNVNS 36
    | | | | | | | | | | | | | | | | | | | |
Db 3 EKKIAKMEKASSVFNVNS 21

RESULT 14
AAR78920
ID AAR78920 standard; peptide; 21 AA.
XX AC AAR78920;
XX DT 27-MAR-1996 (first entry)
XX DE Malaria circumsporozoite 378-398 helper T lymphocyte epitope.
XX KW 378-398; cytotoxic T; CTL; epitope; helper T; HTL; cell;
XX KW lymphocyte; antigens; treatment; disease prevention; tumours;
XX KW viruses; parasites; malaria circumsporozoite; hepatitis B.
XX SQ

```

```

OS Malaria circumsporoite.
XX WO9522317-A1.
XX
XX 24-AUG-1995.
XX
XX 16-FEB-1995; 95WO-US02121.
XX
XX 16-FEB-1994; 94US-0197484.
XX
XX (CYTE-) CYTEL CORP.
XX
XX Ceut RW, Grey H, Sette AD, Vitiello MA;
XX
XX WPI; 1995-302545/39.
XX
XX Compan. inducing cytotoxic T lymphocyte response to pref. viral,
XX bacterial, parasitic or tumour antigens - useful in the treatment
XX and prevention of diseases associated with the antigen e.g.
XX hepatitis B
XX
XX Disclosure: Page 29; 109pp; English.
XX
XX A compsn. which induces a cytotoxic T lymphocyte (CTL) response to
XX an antigen (Ag) in a mammal comprises, a CTL Ag response inducing
XX peptide and a lipid conjugated helper T cell inducing peptide (i.e.
XX AAR78918-R78922). The compsn. induces a CTL response to bacterial,
XX viral or tumour Ags, and is therefore useful in the treatment and
XX prevention of diseases associated with the Ag, e.g. hepatitis B.
XX
XX Sequence 21 AA;
XX
XX Query Match 33.6%; Score 89; DB 16; Length 21;
XX Best Local Similarity 100.0%; Pred. No. 7.7e-05;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 EKXIARMEKASSVFNVNS 36
XX |
XX Db 3 EKXIARMEKASSVFNVNS 21
XX
XX RESULT 15
XX AAR75955
XX ID AAR75955 standard; Peptide: 21 AA.
XX
XX AC AAR75955;
XX
XX 06-MAR-1996 (first entry)
XX
XX P. falciparum CS protein epitope residues 378-398.
XX
XX MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;
XX cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.
XX
XX Plasmodium falciparum.
XX
XX WO9519783-A1.
XX
XX 27-JUL-1995.
XX
XX 25-JAN-1995; 95WO-US01000.
XX
XX 25-JAN-1994; 94US-0186266.
XX
XX (CYTE-) CYTEL CORP.
XX
XX Celis E, Grey HM, Kubo RT, Sette A;
XX
XX WPI; 1995-269270/35.
XX
XX Immunogenic peptide(s) that induce immune response to cancer cells
XX - that express a MAGE-3 protein peptide epitope used in vaccines or
XX adoptive immuno:therapy to induce cytotoxic T lymphocytes
XX
XX Disclosure: Page 14; 44pp; English.
XX
XX AAR75942 is derived from the sequence of the melanoma antigen (MAGE-3)
XX protein and can be used to elicit a primary cytotoxic T lymphocyte
XX response against cells expressing MAGE-3. Synthetic peptides AAR75945-53
XX can be used therapeutically to elicit CTL responses to melanoma, breast,
XX colon, prostate, or other cells which express proteins with this
XX epitope. The peptides have specific HLA-A1 binding capacity. The peptides
XX can be also used in vaccines, esp. combined with peptides such as
XX AAR75955-56, which are T-helper epitopes.
XX
XX Sequence 21 AA;
XX
XX Query Match 33.6%; Score 89; DB 16; Length 21;
XX Best Local Similarity 100.0%; Pred. No. 7.7e-05;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 EKXIARMEKASSVFNVNS 36
XX |
XX Db 3 EKXIARMEKASSVFNVNS 21
XX
XX Search completed: October 10, 2002, 23:10:39
XX Job time : 27 secs

```


GenCore version 5.1.3
 Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 23:09:56 ; Search time 16 Seconds
 (without alignments)
 77.857 Million cell updates/sec

Title: US-09-848-834A-20

Perfect score: 265

Sequence: 1 XHWSYGLRPGSSGSLDEKK.....NVNSSGSPSLHWSYGLRPX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 164566

Minimum DB seq length: 0
 Maximum DB seq length: 51

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127.5	48.1	33	1	US-08-446-692-27
2	127.5	48.1	33	2	US-08-488-351A-27
3	97.5	36.8	49	1	US-08-387-156-4
4	97.5	36.8	49	2	US-08-694-865-4
5	97.5	36.8	49	2	US-08-878-748-4
6	97.5	36.8	49	3	US-09-124-491-4
7	89	33.6	21	1	US-08-186-266-6
8	89	33.6	21	1	US-08-446-692-48
9	89	33.6	21	2	US-08-488-351A-48
10	89	33.6	21	3	US-09-100-409A-54
11	89	33.6	21	5	PCT-US95-02121-97
12	89	33.6	21	5	PCT-US95-13841-20
13	88.5	33.4	40	4	US-09-026-276-35
14	88.5	33.4	40	4	US-09-026-276-34
15	87.5	33.0	44	1	US-07-690-983D-45
16	84.5	31.9	20	1	US-07-690-983D-40
17	84.5	31.9	20	4	US-09-026-276-26
18	84.5	31.9	20	4	US-09-026-276-29
19	84.5	31.9	20	4	US-09-026-276-30
20	84.5	31.9	20	4	US-09-026-276-31
21	84.5	31.9	24	1	US-07-690-983D-43
22	79	29.8	17	5	PCT-US95-02121-96
23	75	28.3	16	2	US-08-817-933A-7
24	73.5	27.7	20	1	US-08-465-167A-20
25	73.5	27.7	20	5	PCT-US92-07218-17
26	71	26.8	17	1	US-08-188-223-6
27	71	26.8	17	4	US-08-968-466-6

28	71	26.8	17	4	US-08-478-546B-6	Sequence 6, Appli
29	70	26.4	15	6	5169933-30	Patent No. 5169933
30	63.5	24.0	16	1	US-08-465-167A-19	Sequence 19, Appl
31	63.5	24.0	16	5	PCT-US92-07218-16	Sequence 16, Appl
32	62	23.4	14	1	US-08-453-588-23	Sequence 23, Appl
33	62	23.4	14	3	US-08-521-079-23	Sequence 23, Appl
34	62	23.4	17	1	US-07-690-983D-18	Sequence 18, Appl
35	62	23.4	18	1	US-07-690-983D-28	Sequence 28, Appl
36	62	23.4	21	1	US-08-305-871A-12	Sequence 12, Appl
37	62	23.4	34	3	US-09-100-414B-82	Sequence 82, Appl
38	62	23.4	34	4	US-09-303-323-82	Sequence 82, Appl
39	61	23.0	16	1	US-08-453-588-26	Sequence 26, Appl
40	61	23.0	16	1	US-08-188-223-2	Sequence 2, Appli
41	61	23.0	16	3	US-08-521-079-26	Sequence 26, Appli
42	61	23.0	16	4	US-08-968-466-2	Sequence 2, Appli
43	61	23.0	16	4	US-08-478-546B-2	Sequence 2, Appli
44	60	22.6	30	3	US-09-100-414B-71	Sequence 71, Appl
45	60	22.6	30	4	US-09-303-323-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1

US-08-446-692-27

; Sequence 27, Application US/08446692

; Patent No. 5759551

; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna

; APPLICANT: Wang, Chang Yi

; APPLICANT: Zamb, Timothy

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin

; REGISTRATION NUMBER: 29,323

; REFERENCE/DOCKET NUMBER: 1151-4146 US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)415-8745

; TELEFAX: (516)751-6849

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 33 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-446-692-27

Query Match 48.1%; Score 127.5; DB 1; Length 33;
 Best Local Similarity 81.8%; Pred No. 6.6e-10;
 Matches 27; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

Qy 18 EKKIARMEKASSVFNVNSSGSPSLHWSYGLRP 50
 |||||
 Db 3 EKKIARMEKASSVFNVNSSGGE---HWSYGLRP 32

```
RESULT 2
US-08-488-351A-27
; Sequence 27, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-27

Query Match 48.18; Score 127.5; DB 2; Length 33;
Best Local Similarity 81.8; Pred. No. 6.6e-10;
Matches 27; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

Qy 18 EKKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
Db 3 EKKIAKMEKASSVFNVNSSG---HWSYGLRP 32

RESULT 3
US-08-387-156-4
; Sequence 4, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
```

```
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-4

Query Match 36.8; Score 97.5; DB 1; Length 49;
Best Local Similarity 42.9; Pred. No. 7.7e-06;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

Qy 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 35

RESULT 4
US-08-694-865-4
; Sequence 4, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
```

Query Match 36.8%; Score 97.5; DB 2; Length 49;

APPLICANT: KUBO, Ralph T.
APPLICANT: GREY, Howard M.
APPLICANT: SETTE, Alessandro
APPLICANT: CELIS, Esteban
TITLE OF INVENTION: INDUCTION OF ANTI-TUMOR CYTOTOXIC
TITLE OF INVENTION: T LYMPHOCYTES IN HUMANS USING
TITLE OF INVENTION: SYNTHETIC PEPTIDE EPITOPES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/186.266
FILING DATE: 25-JAN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/159,339
FILING DATE: 29-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/103,396
FILING DATE: 06-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/027,746
FILING DATE: 05-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,666
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 14137-50-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..21
OTHER INFORMATION: /note= "Plasmodium falciparum CS
OTHER INFORMATION: protein at positions 378-398."
US-08-186-266-6

Query Match 33.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36
|||||
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 8
US-08-446-692-48
Sequence 48, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy

TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446.692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-48

Query Match 33.6%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNS 36
|||||
DB 3 EKKIAKMEKASSVFNVNS 21

RESULT 9
US-08-488-351A-48
Sequence 48, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488.351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446.692
FILING DATE: 7-JUN-1995

```

US-09-100-409A-54
Query Match          33.6%; Score 89; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAMEKASSVENVNS 36
   |||||
Db 3 EKKIAMEKASSVENVNS 21

RESULT 11
PCT-US95-02121-97
; Sequence 97, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Malaria circumsporozoite
; OTHER INFORMATION: 378-398"
PCT-US95-02121-97

Query Match          33.6%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKKIAMEKASSVENVNS 36
   |||||
Db 3 EKKIAMEKASSVENVNS 21

```

RESULT 12
PCT-US95-13841-20
; Sequence 20, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-20
Query Match 33.6%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 EKKIAKMEKASSVFNVNS 36
Db 3 EKKIAKMEKASSVFNVNS 21
RESULT 13
US-09-026-276-35
; Sequence 35, Application US/09026276
; Patent No. 6319503
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, Aprile L
; APPLICANT: Lohnas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: U.S. Patent Application No. 6319503 09\026,276
; CURRENT APPLICATION NUMBER: US/09/026,276
; CURRENT FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 40
; TYPE: PRT

; ORGANISM: Porcine
US-09-026-276-35
Query Match 33.4%; Score 88.5; DB 4; Length 40;
Best Local Similarity 39.2%; Pred. No. 8.5e-05;
Matches 20; Conservative 1; Mismatches 5; Indels 25; Gaps 2;
Qy 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVFNVNSSSG--PSLHWSYGLRP 50
Db 2 HWSYGLRPGQ-----HWSYGLRPGQHWSYGLRP 29
RESULT 14
US-09-026-276-34
; Sequence 34, Application US/09026276
; Patent No. 6319503
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, Aprile L
; APPLICANT: Lohnas, Gerald L
; APPLICANT: Roberts, Steven F
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: U.S. Patent Application No. 6319503 09\026,276
; CURRENT APPLICATION NUMBER: US/09/026,276
; CURRENT FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Porcine
US-09-026-276-34
Query Match 33.4%; Score 88.5; DB 4; Length 41;
Best Local Similarity 39.2%; Pred. No. 8.8e-05;
Matches 20; Conservative 1; Mismatches 5; Indels 25; Gaps 2;
Qy 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVFNVNSSSG--PSLHWSYGLRP 50
Db 2 HWSYGLRPGQ-----HWSYGLRPGQHWSYGLRP 29
RESULT 15
US-07-690-983D-45
; Sequence 45, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.

```
; ; REGISTRATION NUMBER: 29,768
; ; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (202)672-5300
; ; TELEFAX: (202)672-5399
; ; INFORMATION FOR SEQ ID NO: 45:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 44 amino acids
; ; TYPE: amino acid
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: protein
US-07-690-983D-45

Query Match      33.0%; Score 87.5; DB 1; Length 44;
Best Local Similarity 39.2%; Pred. NO. 0.00013;
Matches 20; Conservative 1; Mismatches 5; Indels 25; Gaps 2;

QY  2 HWSYGLRPGSSGSLDEKKIAKMEKASSYFNVVNSSG--PSLHWSYGLRP 50
Db   4 HWSYGLRPG-----HWSYGLRPGEHWSYGLRP 31
```

Search completed: October 10, 2002, 23:13:16
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 2.32906 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-1

Perfect score: 54

Sequence: 1 XHSYGLRPX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	52	96.3	10	1 RHPGG	gonadoliberin - pi
2	52	96.3	10	1 RHSHG	gonadoliberin - sh
3	52	96.3	67	2 I78541	gonadoliberin prec
4	52	96.3	89	2 I51423	gonadoliberin prec
5	52	96.3	90	1 RHMSG	gonadoliberin prec
6	52	96.3	92	1 RHHUG	gonadoliberin prec
7	52	96.3	92	1 RHRTG	gonadoliberin prec
8	48	88.9	10	1 RHAQI	gonadoliberin I -
9	48	88.9	92	2 I50644	gonadoliberin I pr
10	46	85.2	98	2 I50739	gonadotropin-relea
11	43	79.6	374	2 E95361	probable muconate
12	42	77.8	80	1 RHIDL	gonadoliberin I pr
13	42	77.8	91	2 JC7393	medaka-type gonado
14	39	72.2	10	2 A21114	gonadoliberin - ch
15	39	72.2	74	2 I51092	gonadotropin relea
16	39	72.2	82	2 I51180	gonadotropin relea
17	39	72.2	82	2 I51355	gonadotropin relea
18	39	72.2	82	2 I51365	gonadotropin-relea
19	39	72.2	82	2 I51331	gonadotropin relea
20	39	72.2	90	2 JC7395	salmon-type gonado
21	39	72.2	90	2 A23735	gonadoliberin prec
22	39	72.2	90	2 I51095	gonadoliberin prec
23	39	72.2	388	2 T72710	probable fmu prote
24	39	72.2	1444	2 T18856	angiogenesis inhib
25	38	70.4	161	2 D84472	hypothetical prote
26	38	70.4	293	2 G72699	hypothetical prote
27	38	70.4	501	2 T32848	hypothetical prote
28	37	68.5	270	2 F88035	protein M01D1.7 [i
29	37	68.5	345	2 A58519	hypothetical 345 p

30	37	68.5	486	2	T26483	hypothetical prote
31	37	68.5	812	2	T01618	hypothetical prote
32	36	66.7	315	2	A88043	protein C13A10.3 [
33	36	66.7	327	2	E95323	hypothetical prote
34	36	66.7	368	2	T27432	hypothetical prote
35	36	66.7	460	2	D96971	stage V sporulatio
36	36	66.7	517	2	F83573	conserved hypothet
37	36	66.7	927	2	H87568	peptidase, M16 fam
38	35	64.8	82	2	G69446	hypothetical prote
39	35	64.8	300	2	T49933	inorganic pyrophos
40	35	64.8	308	2	D64696	hypothetical prote
41	35	64.8	316	2	A53440	aldose reductase h
42	35	64.8	371	1	E64821	yliI protein precu
43	35	64.8	371	2	E90743	probable dehydroge
44	35	64.8	371	2	A85594	probable dehydroge
45	35	64.8	373	2	AB0604	probable oxidoredu

ALIGNMENTS

RESULT 1

RHPGG

gonadoliberin - piq

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90176; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Note: the synthetic and natural hormones have the same physicochemical and biologic

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 96.3%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0017;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9

Db 2 HWSYGLRP 9

RESULT 2

RHSHG

gonadoliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

A;Residues: 1-10 <BUR>
 A;Note: the natural and synthetic hormones have the same biological activity
 C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
 N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)
 C;Superfamily: gonadoliberin
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0017; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0

QY 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9
 |||||

RESULT 3
 I78541
 gonadoliberin precursor - rhesus macaque (fragment)
 N;Alternate names: luteinizing hormone releasing hormone
 C;Species: Macaca mulatta (rhesus macaque)
 C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
 C;Accession: I78541
 R;Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
 Neuroendocrinology 60, 346-359, 1994
 A;Title: Developmental expression of the genes encoding transforming growth factor alpha and beta in the developing rhesus macaque
 A;Reference number: I58134; MUID:95124501
 A;Accession: I78541
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-67 <RES>
 A;Cross-references: GB:S75918; NID:g912831; PIDN:AAB33096.1; PID:g912832
 C;Superfamily: gonadoliberin

Query Match 96.3%; Score 52; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.013; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0

QY 2 HWSYGLRP 9
 Db 7 HWSYGLRP 14
 |||||

RESULT 4
 I51423
 gonadoliberin precursor - African clawed frog
 N;Alternate names: luteinizing hormone releasing hormone
 C;Species: Xenopus laevis (African clawed frog)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
 C;Accession: I51423
 R;Hayes, W.P.; Wray, S.; Battey, J.F.
 Endocrinology 134, 1835-1845, 1994
 A;Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain
 A;Reference number: I51423; MUID:94185563
 A;Accession: I51423
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-89 <HAY>
 A;Cross-references: GB:L28040; NID:g496291; PIDN:AAA49728.1; PID:g496292
 C;Genetics:
 A;Gene: GnRH-I
 C;Superfamily: gonadoliberin

Query Match 96.3%; Score 52; DB 2; Length 89;
 Best Local Similarity 100.0%; Pred. No. 0.018; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0

QY 2 HWSYGLRP 9
 Db 25 HWSYGLRP 32
 |||||

RESULT 5
 RHMSG
 gonadoliberin precursor - mouse
 N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)
 N;Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
 C;Species: Mus musculus (house mouse)
 C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
 C;Accession: A47578
 R;Wason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolić, S.
 Science 234, 1366-1371, 1986
 A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible for the hypogonadism in the mouse
 A;Reference number: A47578; MUID:87069928
 A;Accession: A47578
 A;Molecule type: DNA
 A;Residues: 1-90 <MAS>
 A;Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
 C;Genetics:
 A;Introns: 45/3; 77/3
 C;Function:
 A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin
 A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
 C;Superfamily: gonadoliberin
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;22-31/Product: gonadoliberin #status predicted <GLB>
 F;35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
 F;22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted <GAP>
 F;31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.018; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0

QY 2 HWSYGLRP 9
 Db 23 HWSYGLRP 30
 |||||

RESULT 6
 RHHUG
 gonadoliberin precursor [validated] - human
 N;Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)
 N;Contains: gonadoliberin-associated protein (GAP); progadoliberin
 C;Species: Homo sapiens (man)
 C;Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
 C;Accession: S05308; A26173; A93342; A90108; A01410; S45718
 R;Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
 Nucleic Acids Res. 17, 6403-6404, 1989
 A;Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone
 A;Reference number: S05308; MUID:89366682
 A;Accession: S05308
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-92 <HAY>
 A;Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
 R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and luteinizing hormone-releasing hormone

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.018; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0

QY 2 HWSYGLRP 9
 Db 23 HWSYGLRP 30
 |||||

RESULT 7
 RHHUG
 gonadoliberin precursor [validated] - human
 N;Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)
 N;Contains: gonadoliberin-associated protein (GAP); progadoliberin
 C;Species: Homo sapiens (man)
 C;Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
 C;Accession: S05308; A26173; A93342; A90108; A01410; S45718
 R;Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
 Nucleic Acids Res. 17, 6403-6404, 1989
 A;Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone
 A;Reference number: S05308; MUID:89366682
 A;Accession: S05308
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-92 <HAY>
 A;Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
 R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and luteinizing hormone-releasing hormone

Query Match 96.3%; Score 52; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.018; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0

QY 2 HWSYGLRP 9
 Db 23 HWSYGLRP 30
 |||||

R: Tan, L.; Rousseau, P.
 Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
 A: Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
 A: Reference number: A90108; MUID: 83126573
 A: Accession: A90108
 A: Molecule type: protein
 A: Residues: 24-33 <TAN>
 A: Experimental source: placental trophoblasts
 R: Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterda
 FEBS Lett. 346, 203-206, 1994
 A: Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
 A: Reference number: S45718; MUID: 94283597
 A: Contents: annotation; degradation pathway of synthetic hormone
 C: Genetics:
 A: Gene: GDB: GNRH; LHRH; GRH
 A: Cross-references: GDB: I33746; OMIM: 227200; OMIM: 152760
 A: Map position: 8p21-8p11.2
 A: Introns: 47/3; 79/3
 C: Function:
 A: Description: gonadoliberein stimulates pituitary secretion of lutropin and follitropin
 C: Superfamily: gonadoliberein
 C: Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
 F: 1-23/Domain: signal sequence #status predicted <SIG>
 F: 24-92/Product: progonadoliberein #status predicted <PGN>
 F: 24-33/Product: gonadoliberein #status predicted <GLN>
 F: 37-92/Product: gonadoliberein-associated protein #status predicted <GAP>
 F: 23/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F: 34/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 96.3%; Score 52; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 25 HWSYGLRP 32
 |||||

RESULT 7
 RHRTG
 gonadoliberein precursor - rat
 N: Alternate names: gonadoliberein-associated protein (GAP); gonadotropin releasing hormon
 N: Contains: gonadoliberein; prolactin release-inhibiting factor
 C: Species: Rattus norvegicus (Norway rat)
 C: Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
 C: Accession: A40147; B26173; A48410
 R: Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
 Mol. Endocrinol. 3, 1257-1262, 1989
 A: Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
 A: Reference number: A40147; MUID: 89384661
 A: Accession: A40147
 A: Molecule type: DNA
 A: Residues: 1-92 <BON>
 A: Cross-references: GB: M31670; NID: g204447; PIDN: AAA41264.1; PID: g204448
 R: Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A: Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
 A: Reference number: A94090; MUID: 86094338
 A: Accession: B26173
 A: Molecule type: mRNA
 A: Residues: 1-92 <ADE>
 A: Cross-references: GB: M12579; NID: g204445; PIDN: AAA41263.1; PID: g204446
 R: Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Bialock, J.E.
 Cell. Mol. Neurobiol. 12, 447-454, 1992
 A: Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone
 A: Reference number: A48410; MUID: 93105480
 A: Accession: A48410
 A: Status: preliminary
 A: Molecule type: mRNA
 A: Residues: 1-92 <MAL>
 A: Cross-references: GB: S50870; NID: g262059; PIDN: AAB24572.1; PID: g262060
 A: Experimental source: thymus

A: Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIP:121083)
 C: Genetics:
 A: Introns: 47/3; 79/3
 C: Function:
 A: Description: stimulates pituitary secretion of lutropin and follitropin
 A: Note: gonadoliberein-associated protein may have prolactin release inhibiting activi
 C: Superfamily: gonadoliberein
 C: Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
 F: 1-23/Domain: signal sequence #status predicted <SIG>
 F: 24-92/Product: progonadoliberein #status predicted <PGN>
 F: 24-33/Product: gonadoliberein #status predicted <GLN>
 F: 37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
 F: 24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
 F: 33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 96.3%; Score 52; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 25 HWSYGLRP 32
 |||||

RESULT 8
 RHAQI
 gonadoliberein I - American alligator
 N: Alternate names: gonadotropin-releasing hormone I
 C: Species: Alligator mississippiensis (American alligator)
 C: Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
 C: Accession: A60066
 R: Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan
 Regul. Pept. 33, 105-116, 1991
 A: Title: Primary structure of two forms of gonadotropin-releasing hormone from brains
 A: Reference number: A60066; MUID: 91352338
 A: Accession: A60066
 A: Molecule type: protein
 A: Residues: 1-10 <LOW>
 C: Superfamily: gonadoliberein
 C: Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F: 1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F: 10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 88.9%; Score 48; DB 1; Length 10;
 Best Local Similarity 87.5%; Pred. No. 0.0087;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
 Db 2 HWSYGLRP 9
 |||||

RESULT 9
 I50644
 gonadoliberein I precursor - chicken
 N: Alternate names: gonadotropin-releasing hormone I
 C: Species: Gallus gallus (chicken)
 C: Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
 R: Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.
 J. Mol. Endocrinol. 11, 19-29, 1993
 A: Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene
 A: Reference number: I50644; MUID: 94059355
 A: Accession: I50644
 A: Status: translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-92 <DU>
 A: Cross-references: EMBL: X69491; NID: g496326; PIDN: CAA49246.1; PID: g311612
 C: Genetics:
 A: Introns: 47/3; 79/3
 C: Superfamily: gonadoliberein

Query Match 88.9%; Score 48; DB 2; Length 92;

Best Local Similarity 87.5%; Pred. No. 0.098;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
|||||:|
Db 25 HWSYGLQP 32

RESULT 10
I50739
gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)
C:Species: Haplochromis burtoni
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I50739
R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.
Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995
A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles
A:Reference number: I50739; MUID:95396797
A:Accession: I50739
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-98 <WHT>
A:Cross-references: EMBL:U31865; NID:g9053398; PIDN:AAC59691.1; PID:g9053399
C:Superfamily: gonadoliberin

Query Match 85.2%; Score 46; DB 2; Length 98;
Best Local Similarity 87.5%; Pred. No. 0.24;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
|||||:|
Db 24 HWSYGLSP 31

RESULT 11
E95361
probable muconate cycloisomerase (EC 5.5.1.1) [imported] - Sinorhizobium meliloti (strain
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: E95361
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
-; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: E95361
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <KUR>
A:Cross-references: GB:AE006459; PIDN:AAK65455.1; PID:g14523923; GSPDB:GNO0165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SWA161
C:Keywords: intramolecular lyase; isomerase

Query Match 79.6%; Score 43; DB 2; Length 374;
Best Local Similarity 85.7%; Pred. No. 3.6;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLR 8
|||||:|
Db 21 HWSYGLR 27

RESULT 12
RH1D1S
gonadoliberin I precursor - sharptooth catfish
N:Alternate names: gonadoliberin, catfish-type; gonadotropin-releasing hormone I (GnRF)
N:Contents: gonadoliberin I; gonadoliberin I-associated protein form I; gonadoliberin I
C:Species: Clarias gariepinus (sharptooth catfish)
C>Date: 30-Sep-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: S45602; S45601; JC1242; S42936; S42937
R:Bogerd, J.; Zandbergen, T.; Andersson, E.; Goos, H.
Eur. J. Biochem. 222, 541-549, 1994
A:Title: Isolation, characterization and expression of cDNAs encoding the catfish-type
A:Reference number: S45600; MUID:94291651
A:Accession: S45602
A:Molecule type: mRNA
A:Residues: 1-80 <BOG1>
A:Cross-references: EMBL:X78049; NID:g459433; PIDN:CAA54971.1; PID:g459434
A:Note: gonadoliberin I-associated protein form I
A:Accession: S45601
A:Molecule type: mRNA
A:Residues: 1-46, 'S', 48-59, 'G', 61-80 <BOG2>
A:Cross-references: EMBL:X78048; NID:g459431; PIDN:CAA54970.1; PID:g459432
A:Note: gonadoliberin I-associated protein form II, presumed to be a polymorphic form
R:Bogerd, J.; Li, K.W.; Janssen-Bommerholt, C.; Goos, H.
Biochem. Biophys. Res. Commun. 187, 127-134, 1992
A:Title: Two gonadotropin-releasing hormones from African catfish (Clarias gariepinus)
A:Reference number: JC1242; MUID:92392313
A:Accession: JC1242
A:Molecule type: protein
A:Residues: 22-31 <BOG3>
A:Experimental source: brain
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin I #status experimental <MAT1>
F:35-80/Product: gonadoliberin I-associated protein #status predicted <MAT2>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi#
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following
Query Match 77.8%; Score 42; DB 1; Length 80;
Best Local Similarity 75.0%; Pred. No. 1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
|||||:|
Db 23 HWSHGLNP 30

RESULT 13
JC7393
medaka-type gonadotropin-releasing hormone precursor - Japanese medaka
C:Species: Oryzias latipes (Japanese medaka)
C>Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7393
R:Okubo, K.; Amano, M.; Yoshiura, Y.; Suetake, H.; Aida, K.
Biochem. Biophys. Res. Commun. 276, 298-303, 2000
A:Title: A novel form of gonadotropin-releasing hormone in the medaka, Oryzias latipes
A:Reference number: JC7393
A:Contents: Brain
A:Accession: JC7393
A:Molecule type: mRNA
A:Residues: 1-91 <OKU>
A:Cross-references: DDBJ:AB041333
C:Comment: This protein plays the roles as a hypophysiotropic factor, and a physiolo#
C:Genetics:
A:Gene: mdgnrh
C:Keywords: brain

Query Match 77.8%; Score 42; DB 2; Length 91;
Best Local Similarity 75.0%; Pred. No. 1.2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9

```
Db      23  HWSFGGLSP 30
||||:| | |
RESULT 14
A21114
gonadoliberin - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C>Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match          72.2%; Score 39; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 0.36;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  HWSYGLRP 9
|||| |
Db      2  HWSYGWLP 9

RESULT 15
I51092
gonadotropin releasing hormone - chinook salmon (fragment)
C:Species: Oncorhynchus tshawytscha (chinook salmon)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Aug-1997
C:Accession: I51092
R:Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.
Mol. Cell. Endocrinol. 84, 167-174, 1992
A:Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.
A:Reference number: I51040; MUID:92267241
A:Accession: I51092
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-74 <KLU>
A:Cross-references: EMBL:X79711; NID:g499322; PID:g499323
C:Genetics:
A:Gene: GnRH
A:Introns: 38/3; 65/3

Query Match          72.2%; Score 39; DB 2; Length 74;
Best Local Similarity 75.0%; Pred. No. 3.2;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  HWSYGLRP 9
|||| |
Db      17 HWSYGWLP 24

Search completed: October 10, 2002, 16:12:00
Job time : 3.40598 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 1.30342 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-1

Perfect score: 54

Sequence: 1 XHWSYGLRPX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	52	96.3	61	GON1_SHEEP	Q28588 ovis aries
2	52	96.3	63	GON1_MESAU	O09163 mesocricetu
3	52	96.3	67	GON1_MACMU	P55247 macaca mula
4	52	96.3	89	GON1_XENLA	P45656 xenopus lae
5	52	96.3	90	GON1_MOUSE	P13562 mus musculus
6	52	96.3	91	GON1_PIG	P49921 sus scrofa
7	52	96.3	92	GON1_HUMAN	P01148 homo sapien
8	52	96.3	92	GON1_RAT	P07490 rattus norv
9	52	96.3	92	GON1_TUPGB	Q95335 tupaia glis
10	48	88.9	10	GON1_ALLMI	P37041 alligator m
11	48	88.9	92	GON1_CHICK	P37042 gallus gall
12	46	85.2	94	GON1_HAPBU	P51918 haplochromi
13	46	85.2	95	GON1_MORSA	O73812 morone saxa
14	46	85.2	95	GON1_PAGMA	P70074 pagrus majo
15	46	85.2	95	GON1_SPAAU	P51919 sparus aura
16	46	85.2	99	GON1_DICLA	Q91a10 dicentrarch
17	44	81.5	90	GON8_RANDY	Q91a02 rana dybows
18	43	79.6	92	GON1_CAVPO	O54713 cavia porce
19	42	77.8	80	GON1_CLAGA	P33439 clarias gar
20	41	75.9	10	GON1_CLUPA	P81749 clupea pall
21	41	75.9	110	YHBT_ACTAC	P96769 actinobacil
22	39	72.2	10	GON3_ONCKE	P20367 oncorhynch
23	39	72.2	74	GON3_ONCMY	P55246 oncorhynch
24	39	72.2	74	GON3_ONCTS	Q92097 oncorhynch
25	39	72.2	82	GON3_ONCMA	P30973 oncorhynch
26	39	72.2	82	GON3_SALSA	P35629 salmo salar
27	39	72.2	82	GON3_SALTR	P45653 salmo trutt
28	39	72.2	89	GON3_PORNO	P51922 porichthys
29	39	72.2	90	GON3_DICLA	Q91a09 dicentrarch
30	39	72.2	90	GON3_HAPBU	P45652 haplochromi
31	39	72.2	90	GON3_PAGMA	P51921 pagrus majo
32	39	72.2	90	GON3_SPAAU	P51923 sparus aura
33	39	72.2	94	GON3_CARAU	P51917 carassius a

ALIGNMENTS

RESULT 1

ID	GON1_SHEEP	STANDARD;	PRT;	61 AA.
AC	Q28588;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)			
DE	(Luteinizing hormone releasing hormone I) (Gonadotropin releasing			
DE	hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]			
DE	(Fragment).			
GN	GNRH1 OR GNRH OR LHRH.			
OS	Ovis aries (Sheep).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
OX	NCBI_TaxID=9940;			
RN	[1]			
RP	SEQUENCE OF 12-61 FROM N.A.			
RC	STRAIN-WESTERN RANGE; TISSUE-Hypothalamus;			
RA	Rodriguez R.E., Wise M.E.;			
RL	Submitted (Oct-1993) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE OF 1-10.			
RA	MEDLINE=72094314; PubMed=4550508;			
RA	Burgus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,			
RA	Fellows R., Blackwell R., Vale W., Gullermin R.;			
RT	"Primary structure of the ovine hypothalamic luteinizing hormone-			
RT	releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass			
RT	spectrometry-decapeptide-Edman degradation).";			
RL	Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).			
CC	-!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS: IT STIMULATES			
CC	THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING			
CC	HORMONES.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U02517; AAA03433.1; -			
DR	PIR; A93780; RHSHG.			
DR	InterPro; IPR002012; GNRH.			
DR	Pfam; PF00446; GNRH; 1.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;			
KW	Placenta.			
FT	NON_TER	1	1	
FT	CHAIN	1	>61	PROGONADOLIBERIN I.
FT	PEPTIDE	1	10	GONADOLIBERIN I.
FT	PEPTIDE	14	>61	GNRH-ASSOCIATED PEPTIDE I.

Q92106 rutillus rut
O28696 archaeglob
P45377 mus musculus
P75804 escherichia
O55197 rattus norv
O09047 mus musculus
P25472 clostridium
P37043 gallus gall
P27429 squalus aca
Q91a08 dicentrarch
P37044 haplochromi
O73811 morone saxa

	Matches	8;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	2	HWSYGRLP	9							
Db	2	HWSYGRLP	9							
RESULT 3										
ID	GONI_MACMU	STANDARD;	PRT;	67	AA.					
AC	P55247;									
DT	01-OCT-1996	(Rel. 34, Created)								
DT	01-OCT-1996	(Rel. 34, Last sequence update)								
DT	30-MAY-2000	(Rel. 39, Last annotation update)								
DE	Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I) luteinizing hormone releasing hormone I] (Gonadotropin releasing hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I]									
DE	(Fragment).									
DE	GNRH1 OR GNRH OR LHRH.									
OS	Macaca mulatta (Rhesus macaque).									
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
OC	Mammalia; Euthera; Primates; Catarrhini; Cercopithecoidea;									
OC	Cercopithecinae; Macaca.									
OX	NCBI_TaxID=9544;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	TISSUE-Hypothalamus;									
RX	MEDLINE=95124501; PubMed=7545971;									
RA	Ma Y.J., Costa M.E., Ojeda S.R.;									
RT	"developmental expression of the genes encoding transforming growth factor alpha and its receptor in the hypothalamus of female rhesus macaques.";									
RT	RT									
RL	Neuroendocrinology 60:346-359(1994).									
CC	-I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.									
CC	-I- SUBCELLULAR LOCATION: Secreted.									
CC	-I- SIMILARITY: BELONGS TO THE GNRH FAMILY.									
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to licensel@isb-sib.ch).									
DR	EMBL: S75918: AAB33096.1;									
DR	InterPro: IPR002012; GnRH.									
DR	Pfam: PF00446; GnRH; 1.									
DR	PROSITE: PS00473; GNRH; 1.									
KW	Cleaveage on pair of basic residues; Hormone; Amidation; Hypothalamus; Signal.									
FT	NON_TER	1	1							
FT	SIGNAL	<1	5	BY SIMILARITY.						
FT	CHAIN	6	>67	PROGONADOLIBERIN I.						
FT	PEPTIDE	6	15	GONADOLIBERIN I.						
FT	PEPTIDE	19	>67	GNRH-ASSOCIATED PEPTIDE I.						
FT	ACT_SITE	8	8	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY (BY SIMILARITY).						
FT	MOD_RES	6	6	PYROLIDONE CARBOXYLIC ACID (BY SIMILARITY).						
FT	MOD_RES	15	15	AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY SIMILARITY).						
FT	NON_TER	67	67							
SQ	SEQUENCE	67 AA;	7573 MW;	505394DAA261A3F2	CRC64;					
Query Match 96.3%; Score 52; DB 1; Length 67;										
Best Local Similarity 100.0%; Pred. No. 0.0036;										
Matches	8;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	2	HWSYGRLP	9							


```

Db      7 HWSYGLRP 14
RESULT 4
GONL_XENLA
ID      GONL_XENLA  STANDARD;      PRT;      89 AA.
AC      P4566;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
DE      (LH-RH) (Luliberin I).
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Forebrain;
RX      MEDLINE=94185563; PubMed=8137750;
RA      Hayes W.P., Wray S., Battey J.F.;
RT      "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT      mammalian-like expression pattern and conserved domains in
RT      GNRH-associated peptide, but brain onset is delayed until
RT      metamorphosis."
RL      Endocrinology 134:1835-1844(1994).
CC      -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L28040; AAA49728.1; -.
DR      InterPro; IPR002012; GNRH.
DR      Pfam; PF00446; GNRH; 1.
DR      PRINTS; PR01541; GONADOLIBRN1.
DR      PROSITE; PS00473; GNRH; 1.
KW      Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW      Signal.
FT      SIGNAL          1      23
FT      CHAIN           24      89
FT      PEPTIDE         24      33
FT      PEPTIDE         37      89
FT      PEPTIDE         37      85
FT      MOD_RES         24      24
FT      MOD_RES         33      33
FT      MOD_RES         33      33
SQ      SEQUENCE      89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;

Query Match          96.3%; Score 52; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HWSYGLRP 9
Db      25 HWSYGLRP 32
RESULT 5
GONL_MOUSE
ID      GONL_MOUSE  STANDARD;      PRT;      90 AA.
AC      P3562;
DT      01-JAN-1990 (Rel. 13, Created).
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)

```

```

DE      Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE      (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE      hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor
DE      I].
GN      GNRH1 OR GNRH.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87069928; PubMed=3024317;
RA      Mason A.J., Haylick J.S., Zoeller R.T., Young W.S. III,
RA      Phillips H.S., Nikolics K., Seeburg P.H.;
RT      "A deletion truncating the gonadotropin-releasing hormone gene is
RT      responsible for hypogonadism in the hpg mouse."
RL      Science 234:1366-1371(1986).
CC      -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC      THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC      HORMONES.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M14872; AAA37717.1; -.
DR      MGD; MGI:95789; Gnrh.
DR      InterPro; IPR002012; GNRH.
DR      InterPro; IPR004079; GonadoliberinI.
DR      Pfam; PF00446; GNRH; 1.
DR      PRINTS; PR01541; GONADOLIBRN1.
DR      PROSITE; PS00473; GNRH; 1.
KW      Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW      Placenta; Signal.
FT      SIGNAL          1      21
FT      CHAIN           22      90
FT      PEPTIDE         22      31
FT      PEPTIDE         35      90
FT      ACT_SITE        24      24
FT      MOD_RES         22      22
FT      MOD_RES         31      31
FT      SEQUENCE        90 AA; 10337 MW; 1C0766FA482684D9 CRC64;

Query Match          96.3%; Score 52; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HWSYGLRP 9
Db      23 HWSYGLRP 30
RESULT 6
GONL_PIG
ID      GONL_PIG    STANDARD;      PRT;      91 AA.
AC      P49921;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE      (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE      hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].
GN      GNRH1 OR GNRH.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```

OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RA Weesner G.D., Matteri R.L., Becker B.A.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=72114303; PubMed=4946067;
 RA Baba Y., Matsuo H., Schally A.V.;
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.
 RT Confirmation of the proposed structure by conventional sequential
 RT analyses.";
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
 RN [3]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72065376; PubMed=4942726;
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 RT phase method.";
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
 RN [4]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72117544; PubMed=4946275;
 RA Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L32864; AAA31066.1; -
 CC PIR: A01411; RHPGG.
 CC InterPro: IPR002012; GNRH.
 CC DR InterPro: IPR004079; GonadoliberinI.
 CC DR Pfam: PF00446; GNRH; 1.
 CC DR PRINTS: PR01541; GONADOLIBRNI.
 CC DR PROSITE: PS00473; GNRH; 1.
 CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 CC Placenta; Signal.
 CC
 CC FT SIGNAL 1 23 PROGONADOLIBERIN I.
 CC FT CHAIN 24 91 GONADOLIBERIN I.
 CC FT PEPTIDE 24 33 GNRH-ASSOCIATED PEPTIDE I.
 CC FT PEPTIDE 34 91 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 CC FT ACT_SITE 26 26 ACTIVITY.
 CC FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 CC SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
 Query Match 96.3%; Score 52; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.0048;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSVGLRP 9
 Db 25 HWSVGLRP 32
 |||||
 ?
 RESULT 7
 GON1_HUMAN STANDARD; PRT; 92 AA.
 AC P01148;

DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated
 DE peptide I].
 GN GNRH1 OR GNRH OR LHRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89366682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 RT hormone gene.";
 RL Nucleic Acids Res. 17:6403-6403(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85012739; PubMed=6090951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 RT releasing hormone.";
 RL Nature 311:666-668(1984).
 RN [4]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=83126573; PubMed=6760865;
 RA Tan L., Rousseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide
 RT biosynthesized in the human placenta.";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 CC Lutrepulse or Lutrel (Ferring Pharmaceuticals) and Relisorm
 CC (Serono).
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X01059; CAA25526.1; -
 CC DR EMBL: M12578; AAA35916.1; -
 CC DR EMBL: X15215; CAA33285.1; -
 CC DR PIR: A01410; RHHUG.
 CC DR PIR: A26173; A26173.
 CC DR PIR: S05308; S05308.
 CC DR MIM: 152760; -
 CC InterPro: IPR002012; GNRH.
 CC DR InterPro: IPR004079; GonadoliberinI.
 CC DR Pfam: PF00446; GNRH; 1.
 CC DR PRINTS: PR01541; GONADOLIBRNI.
 CC DR PROSITE: PS00473; GNRH; 1.
 CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 CC Placenta; Pharmaceutical; Signal.
 CC SIGNAL 1 23

```

FT CHAIN      24 92  PROGNADOLIBERIN I.
FT PEPTIDE    24 33  GONADOLIBERIN I.
FT PEPTIDE    37 92  GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE   26 26  APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT            26 26  ACTIVITY.
FT MOD_RES    24 24  PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES    33 33  AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT CONFLICT   16 16  W -> S (IN REF. 3).
FT SEQUENCE   92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match      96.3%; Score 52; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
Db 25 HWSYGLRP 32

RESULT 8
GONI_RAT
ID GONI_RAT STANDARD; PRT; 92 AA.
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); Prolactin release-inhibiting factor
DE I].
DE GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RA "Isolation of the gene and hypothalamic cDNA for the common precursor
RA of gonadotropin-releasing hormone and prolactin release-inhibiting
RA factor in human and rat.";
RA Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Thymus;
RC MEDLINE=93105480; PubMed=1468115;
RX Maier C.C., Marchetti B., Leboeuf R.D., Bialock J.E.;
RA "Thymocytes express a mRNA that is identical to hypothalamic
RA luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RN SEQUENCE OF 1-47 FROM N.A.
RP TISSUE=Heart;
RC MEDLINE=87149087; PubMed=3547652;
RX Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RA "Two mammalian genes transcribed from opposite strands of the same
RA DNA locus.";
RL Science 235:1514-1517(1987).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
-----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S08070; AAB24572.1; -
DR EMBL; M12579; AAA41263.1; -
DR EMBL; M31670; AAA41264.1; -
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -
DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR PIR; B26173; RHRTG.
DR PIR; A48410; A48410.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL      1 23  PROGNADOLIBERIN I.
FT CHAIN       24 92  GONADOLIBERIN I.
FT PEPTIDE     24 33  PROLACTIN RELEASE-INHIBITING FACTOR I.
FT PEPTIDE     37 92  APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE    26 26  ACTIVITY.
FT MOD_RES     24 24  PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES     33 33  AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT SEQUENCE    92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match      96.3%; Score 52; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRP 9
Db 25 HWSYGLRP 32

RESULT 9
GONI_TUPGB
ID GONI_TUPGB STANDARD; PRT; 92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); GNRH-associated peptide I].
DE GNRH1 OR GNRH.
OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.
OX NCBI_TaxID=9396;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Hypothalamus;
RC MEDLINE=97079639; PubMed=8921350;
RX Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RA "Characterization of two new preproGnrh mRNAs in the tree shrew:
RA first direct evidence for mesencephalic Gnrh gene expression in a
RA placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
-----

```



```
FT PEPTIDE 36 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP)
FT (BY SIMILARITY).
SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match 85.2%; Score 46; DB 1; Length 95;
Best Local Similarity 87.5%; Pred. No. 0.065;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | |
Db 24 HWSYGLSP 31

RESULT 14
GONI_PAGMA STANDARD; PRT; 95 AA.
AC P70074;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
DE (LH-RH I) (Luliberin I).
GN GNRH1.
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Pagrus.
OX NCBI_TaxID=143350;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95268499; PubMed=7749463;
RA Gonthif Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-
releasing hormone from the gilthead seabream (Sparus aurata).";
RL Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE=Brain;
RX MEDLINE=95083645; PubMed=7991588;
RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
brains of one species.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D86582; BAA13129.1; -.
CC InterPro; IPR002012; GNRH.
CC Pfam; PF00446; GNRH; 1.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 24 95
FT PEPTIDE 24 33 POTENTIAL.
FT MOD_RES 37 95 GONADOLIBERIN I.
FT MOD_RES 24 24 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT PYRROLIDONE CARBOXYLIC ACID (BY
SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP)
FT (BY SIMILARITY).
SQ SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;

Query Match 85.2%; Score 46; DB 1; Length 95;
Best Local Similarity 87.5%; Pred. No. 0.065;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | |
Db 25 HWSYGLSP 32
```

```
RESULT 15
GONI_SPAU STANDARD; PRT; 95 AA.
AC P51919;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
DE (LH-RH I) (Luliberin I) (SBGNRH).
GN GNRH1.
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95268499; PubMed=7749463;
RA Gonthif Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-
releasing hormone from the gilthead seabream (Sparus aurata).";
RL Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE=Brain;
RX MEDLINE=95083645; PubMed=7991588;
RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
brains of one species.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U30320; AAA75469.1; -.
CC InterPro; IPR002012; GNRH.
CC Pfam; PF00446; GNRH; 1.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 25
FT CHAIN 26 95
FT PEPTIDE 26 35 GONADOLIBERIN I.
FT MOD_RES 39 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 35 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
SQ SEQUENCE 95 AA; 10753 MW; 49313FD6FD6B87DA CRC64;

Query Match 85.2%; Score 46; DB 1; Length 95;
Best Local Similarity 87.5%; Pred. No. 0.065;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | |
Db 27 HWSYGLSP 34

Search completed: October 10, 2002, 16:06:26
Job time : 2.30342 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 3.88889 Seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-1
Perfect score: 54
Sequence: 1 XHSYGLRXPX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SPTRMBL19.*
 - 2: sp_archaea.*
 - 3: sp_bacteria.*
 - 4: sp_fungi.*
 - 5: sp_human.*
 - 6: sp_invertebrate.*
 - 7: sp_mammal.*
 - 8: sp_mhc.*
 - 9: sp_organelle.*
 - 10: sp_phage.*
 - 11: sp_plant.*
 - 12: sp_rodent.*
 - 13: sp_virus.*
 - 14: sp_vertebrate.*
 - 15: sp_unclassified.*
 - 16: sp_virus.*
 - 17: sp_bacteriophage.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	96.3	90	13 Q90Y63	Q90Y63 rana catesb
2	52	96.3	91	13 Q9PRH0	Q9PRH0 anguilla ja
3	46	85.2	87	13 Q9YI26	Q9YI26 sparus aura
4	43	79.6	374	16 Q92YR6	Q92YR6 rhizobium m
5	42	77.8	91	13 Q9DGC8	Q9DGC8 oryzias lat
6	40	74.1	367	16 Q987W1	Q987W1 rhizobium l
7	39	72.2	33	13 Q9PT34	Q9PT34 oncorhynch
8	39	72.2	33	13 Q9W7G0	Q9W7G0 oncorhynch
9	39	72.2	54	13 Q9W0W9	Q9W0W9 oncorhynch
10	39	72.2	62	13 Q90ZE1	Q90ZE1 oncorhynch
11	39	72.2	82	13 Q918Q0	Q918Q0 oncorhynch
12	39	72.2	82	13 Q918P9	Q918P9 oncorhynch
13	39	72.2	82	13 Q92094	Q92094 oncorhynch
14	39	72.2	82	13 Q9W7G1	Q9W7G1 oncorhynch
15	39	72.2	82	13 Q90VY3	Q90VY3 oncorhynch
16	39	72.2	88	13 Q9PSY9	Q9PSY9 sparus aura

17	39	72.2	90	13 Q9DD49	Q9DD49 oryzias lat
18	39	72.2	94	13 Q9DDDB	Q9DDDB brachydanio
19	39	72.2	94	13 Q9DEH6	Q9DEH6 carassius a
20	39	72.2	94	13 Q9DEH5	Q9DEH5 carassius a
21	39	72.2	388	17 Q9YD14	Q9YD14 aeropyrum p
22	39	72.2	1444	5 Q17591	Q17591 caenorhabdi
23	38	70.4	161	10 Q9ZUC1	Q9ZUC1 arabidopsis
24	38	70.4	205	10 Q9AWR9	Q9AWR9 oryza sativ
25	38	70.4	208	5 Q9VM18	Q9VM18 drosophila
26	38	70.4	236	10 Q9FYR4	Q9FYR4 arabidopsis
27	38	70.4	293	17 Q9YD98	Q9YD98 aeropyrum p
28	38	70.4	334	5 Q44865	Q44865 caenorhabdi
29	38	70.4	1687	3 Q9C024	Q9C024 schizosacch
30	37	68.5	218	2 Q9KXJ2	Q9KXJ2 streptomyce
31	37	68.5	270	5 Q17228	Q17228 caenorhabdi
32	37	68.5	322	2 Q9F3C9	Q9F3C9 streptomyce
33	37	68.5	486	5 Q45910	Q45910 caenorhabdi
34	37	68.5	812	10 Q64620	Q64620 arabidopsis
35	37	68.5	1660	2 Q9EVR7	Q9EVR7 xenorhabdus
36	37	68.5	2091	3 P78616	P78616 emericella
37	36	66.7	89	2 Q9ZNI3	Q9ZNI3 pseudomonas
38	36	66.7	306	10 Q9MSQ3	Q9MSQ3 petunia hyb
39	36	66.7	315	5 P91045	P91045 caenorhabdi
40	36	66.7	327	16 Q92ZJ2	Q92ZJ2 rhizobium m
41	36	66.7	368	5 Q9U1R7	Q9U1R7 caenorhabdi
42	36	66.7	460	16 Q97L10	Q97L10 clostridium
43	36	66.7	517	16 Q915V1	Q915V1 pseudomonas
44	36	66.7	596	11 P97406	P97406 mus musculu
45	36	66.7	596	11 Q91UZ6	Q91UZ6 mus musculu

ALIGNMENTS

RESULT 1

Q90Y63 PRELIMINARY; PRT; 90 AA.
AC Q90Y63;
DT 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE GONADOTROPIN-RELEASING HORMONE.
GN GNRH.

OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
ON NCBI_TaxID=8400;
RX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21102951; PubMed=11170016;
RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GnrH1 and GnrH2
precursors from bullfrog (Rana catesbeiana).";
RL J. Exp Zool. 289:190-201(2001).
DR EMBL; AF188754; AAL05972.1; -;
SQ SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;

Query Match 96.3%; Score 52; DB 13; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 26 HWSYGLRP 33

RESULT 2

Q9PRH0 PRELIMINARY; PRT; 91 AA.
AC Q9PRH0;
DT 01-MAY-2000 (TREMREL. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE GONADOLIBERIN PRECURSOR (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-
 DE RH) (LULIBERIN).
 OS Anguilla japonica (Japanese eel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguilloidei;
 OC Anguillidae; Anguilla.
 OX NCBI_TaxID=7937;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=BRAIN;
 RA Okubo K., Suetake H., Aida K.;
 RC "Expression of two gonadotropin-releasing hormone (GNRH) precursor
 RT genes in various tissues of the Japanese eel and evolution of GNRH.";
 RL Zool. Sci. 16:471-478(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Okubo K., Suetake H., Aida K.;
 RT "A splicing variant for the prepro-mammalian gonadotropin-releasing
 RT hormone (prepro-mGNRH) mRNA is present in the brain and various
 RT peripheral tissues of the Japanese eel.";
 RL Zool. Sci. 16:645-651(1999).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC EMBL; AB026989; BAA82608.1; -.
 CC EMBL; AB026991; BAA83597.1; -.
 DR InterPro; IPR002012; GNRH.
 DR InterPro; IPR004079; Gonadoliberin1.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBRN1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Amidation; Hormone; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 32 MGNRH.
 FT CHAIN 33 91 GNRH ASSOCIATED PEPTIDE.
 SQ SEQUENCE 91 AA; 9893 MW; BA15C9DC08434A7B CRC64;
 Query Match 96.3%; Score 52; DB 13; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HWSYGLRP 9
 Db 24 HWSYGLRP 31
 RESULT 3
 ID Q9YI26 PRELIMINARY; PRT; 87 AA.
 AC Q9YI26;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
 DE (LULIBERIN) (FRAGMENT).
 OS Sparus aurata (Gilthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorphi; Acanthopterygii; Perciformes; Percoidae;
 OC Sparidae; Sparus.
 OX NCBI_TaxID=8175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=OVARY;
 RA Nabissi M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR EMBL; AF046801; AAD02427.1; -.
 DR InterPro; IPR002012; GNRH.

DR Pfam; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Amidation; Hormone.
 FT NON_TER 1 87
 FT NON_TER 87 87
 SQ SEQUENCE 87 AA; 9871 MW; OD2463533D96782A CRC64;
 Query Match 85.2%; Score 46; DB 13; Length 87;
 Best Local Similarity 87.5%; Pred. No. 0.25;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 HWSYGLRP 9
 Db 22 HWSYGLRP 29
 RESULT 4
 ID Q92YR6 PRELIMINARY; PRT; 374 AA.
 AC Q92YR6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PUTATIVE MUONATE CYCLOISOMERASE (EC 5.5.1.1).
 GN SMA1461.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSyma (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RC MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barloy-Hudler F., Bowser L., Capela D., Galibert F., Gouzy J.,
 RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSyma megaplasmid";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 DR EMBL; AE007266; AAK65455.1; -.
 KW Isomerase; Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 374 AA; 40999 MW; 0369AA67790B23D6 CRC64;
 Query Match 79.6%; Score 43; DB 16; Length 374;
 Best Local Similarity 85.7%; Pred. No. 4.4;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 HWSYGLRP 8
 Db 21 HWSYGIR 27
 RESULT 5
 ID Q9DGC8 PRELIMINARY; PRT; 91 AA.
 AC Q9DGC8;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PREPRO-GONADOTROPIN-RELEASING HORMONE.
 GN MDGNRH.
 OS Oryzias latipes (Medaka fish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorphi; Acanthopterygii; Perciformes; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=BRAIN;

```

RX MEDLINE=20462954; PubMed=11006121;
RA Okubo K., Anano M., Yoshiura Y., Suetake H., Aida K.;
RT "A Novel Form of Gonadotropin-Releasing Hormone in the Medaka, Oryzias latipes.";
RL Biochem. Biophys. Res. Commun. 276:298-303(2000).
DR EMBL; AB041333; BAB16303.1; -.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF004446; GnRH; 1.
DR PROSITE; PS00473; GnRH; UNKNOWN_1.
FT CHAIN 22 31
SQ SEQUENCE 91 AA; 10307 MW; A00F2BED6FD6E0B5 CRC64;

Query Match 77.8%; Score 42; DB 13; Length 91;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 23 HWSFGLSP 30

RESULT 6
Q987W1 PRELIMINARY; PRT; 367 AA.
AC Q987W1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE MLL6883 PROTEIN.
GN MLL6883.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003010; BAB53089.1; -.
KW Complete proteome.
SQ SEQUENCE 367 AA; 41874 MW; 30FA5D0FA44ABE7F CRC64;

Query Match 74.1%; Score 40; DB 16; Length 367;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGL 7
Db 352 HWSFGL 357

RESULT 7
Q9PT34 PRELIMINARY; PRT; 33 AA.
AC Q9PT34;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GNADOLIBERIN (GNADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
GN GNRH1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

RX MEDLINE=20462954; PubMed=11006121;
RA Okubo K., Anano M., Yoshiura Y., Suetake H., Aida K.;
RT "A Novel Form of Gonadotropin-Releasing Hormone in the Medaka, Oryzias latipes.";
RL Biochem. Biophys. Res. Commun. 276:298-303(2000).
DR EMBL; AB041333; BAB16303.1; -.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF004446; GnRH; 1.
DR PROSITE; PS00473; GnRH; UNKNOWN_1.
FT CHAIN 22 31
SQ SEQUENCE 91 AA; 10307 MW; A00F2BED6FD6E0B5 CRC64;

Query Match 77.8%; Score 42; DB 13; Length 91;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 23 HWSFGLSP 30

RESULT 6
Q987W1 PRELIMINARY; PRT; 367 AA.
AC Q987W1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE MLL6883 PROTEIN.
GN MLL6883.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003010; BAB53089.1; -.
KW Complete proteome.
SQ SEQUENCE 367 AA; 41874 MW; 30FA5D0FA44ABE7F CRC64;

Query Match 74.1%; Score 40; DB 16; Length 367;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGL 7
Db 352 HWSFGL 357

RESULT 7
Q9PT34 PRELIMINARY; PRT; 33 AA.
AC Q9PT34;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GNADOLIBERIN (GNADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
GN GNRH1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
```

```

OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
RT differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110533; AAD43461.1; -.
DR InterPro; IPR002047; AKH.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00256; AKH; UNKNOWN_1.
DR PROSITE; PS00473; GnRH; 1.
KW Amulation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;

Query Match 72.2%; Score 39; DB 13; Length 33;
Best Local Similarity 75.0%; Pred. No. 1.7;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 25 HWSFGLSP 32

RESULT 8
Q9W7G0 PRELIMINARY; PRT; 33 AA.
AC Q9W7G0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GNADOLIBERIN (GNADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
GN GNRH2.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
RT differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110993; AAD43463.1; -.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Amulation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;

Query Match 72.2%; Score 39; DB 13; Length 33;
Best Local Similarity 75.0%; Pred. No. 1.7;
```

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
||||| 1
Db 25 HWSYGLWP 32

RESULT 9

Q90W09 ID Q90W09 PRELIMINARY; PRT; 54 AA.
AC Q90W09;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GONADOTROPIN-RELEASING HORMONE PRECURSOR II ISOFORM D
DE (GONADOTROPIN-RELEASING HORMONE PRECURSOR II ISOFORM C).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=21232987; PubMed=11335940;
RA Uzbekova S., Ferriere F., Guiguen Y., Bailhache T., Breton B.,
RA Lareyre J.J.;
RT "Stage-dependent and alternative splicing of sGnRH messengers in
rainbow trout testis during spermatogenesis.";
RL Mol. Reprod. Dev. 59:1-10(2001).
DR EMBL; AF269107; AAK54679.1; -.
DR EMBL; AF269106; AAK54678.1; -.
SQ SEQUENCE 54 AA; 5963 MW; 06BF365F658E08B7 CRC64;

Query Match 72.2%; Score 39; DB 13; Length 54;

Best Local Similarity 75.0%; Pred. No. 3;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
||||| 1
Db 25 HWSYGLWP 32

RESULT 10

Q90ZE1 ID Q90ZE1 PRELIMINARY; PRT; 62 AA.
AC Q90ZE1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PRE-PRO-SGHRH-I PROTEIN (FRAGMENT).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Uzbekova S., Ferriere F., Breton B.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF231728; AAK82957.1; -.
FT NON_TER 1 1
FT 62 62
SQ SEQUENCE 62 AA; 7018 MW; E388AAA57E96B8DC CRC64;

Query Match 72.2%; Score 39; DB 13; Length 62;

Best Local Similarity 75.0%; Pred. No. 3.4;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
||||| 1
Db 7 HWSYGLWP 14

RESULT 11

Q918Q0 ID Q918Q0 PRELIMINARY; PRT; 82 AA.
AC Q918Q0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Ferriere F., Bailhache T., Jégo P.;
RT "Oncorhynchus mykiss sGnRH-I cDNA from brain.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF232212; AAF91280.1; -.
DR InterPro; IPR002047; AKH.
DR InterPro; IPR002012; GnrH.
DR Pfam; PF00446; GnrH; 1.
DR PROSITE; PS00256; AKH; UNKNOWN_1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
SQ SEQUENCE 82 AA; 9198 MW; 7595A0B96556A69 CRC64;

Query Match 72.2%; Score 39; DB 13; Length 82;

Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
||||| 1
Db 25 HWSYGLWP 32

RESULT 12

Q918P9 ID Q918P9 PRELIMINARY; PRT; 82 AA.
AC Q918P9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Ferriere F., Bailhache T., Jégo P.;
RT "Oncorhynchus mykiss sGnRH-II cDNA in the brain.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF232213; AAF91281.1; -.
DR InterPro; IPR002012; GnrH.
DR Pfam; PF00446; GnrH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
SQ SEQUENCE 82 AA; 9203 MW; 8053F4F221A0FF08 CRC64;

Query Match 72.2%; Score 39; DB 13; Length 82;

RA Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring.";
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR PIR: B60066; RHAQ2.
 DR PIR: A61126; A61126.
 DR PIR: B46030; B46030.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;
 Query Match 21.2%; Score 34; DB 1; Length 10;
 Best Local Similarity 62.5%; Pred. No. 46;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 23 HWSYGLRP 30
 Db 2 HWSHGWP 9
 RESULT 5
 GONL_SQUAC STANDARD; PRT; 10 AA.
 ID GONL_SQUAC
 AC P27A29;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
 DE (Luliberin).
 OS Squalus acanthias (Spiny dogfish).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 CC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
 OX NCBI_TaxID=7797;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=92333300; PubMed=1631133;
 RA Lovejoy D.A., Fischer W.H., Ngamwongchon S., Craig A.G.,
 RA Nahornlak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
 RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
 RT dogfish brain provides insight into GNRH evolution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;
 Query Match 21.2%; Score 34; DB 1; Length 10;
 Best Local Similarity 62.5%; Pred. No. 46;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 23 HWSYGLRP 30
 Db 2 HWSHGWP 9
 RESULT 6
 PCW4_PACGO STANDARD; PRT; 26 AA.
 ID PCW4_PACGO
 AC P82426;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Ponericin W4.
 OS Pachycondyla goeldii (Ponerine ant).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 CC Formicidae; Ponerinae; Pachycondyla.
 OX NCBI_TaxID=118888;
 RN [1]
 RP SEQUENCE, AND FUNCTION.
 RC TISSUE=Venom;
 RX PubMed=11279030;
 RA Orivel J., Redeker V., Le Caer J.-P., Krier F., Revol-Junelles A.-M.,
 RA Longeon A., Chafotte A., Dejean A., Rossier J.;
 RT "Ponericins, new antibacterial and insecticidal peptides from the
 RT venom of the ant Pachycondyla goeldii.";
 RL J. Biol. Chem. 276:17823-17829(2001).
 CC -!- FUNCTION: BROAD SPECTRUM OF ACTIVITY AGAINST BOTH GRAM-POSITIVE
 CC AND GRAM-NEGATIVE BACTERIA AND S.CEREVISIAE. HAS INSECTICIDAL
 CC AND HEMOLYTIC ACTIVITIES.
 CC -!- MASS SPECTROMETRY: MW=2851.81; METHOD=MALDI.
 KW Antibiotic; Insect immunity; Fungicide; Hemolysis.
 SQ SEQUENCE 26 AA; 2853 MW; E361FBB1F5D4F70F CRC64;
 Query Match 20.0%; Score 32; DB 1; Length 26;
 Best Local Similarity 40.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 16 GVEGPSLHWSYGLRP 30
 Db 1 GIWGTALKMGVKLLP 15
 RESULT 7
 GON3_PETMA STANDARD; PRT; 10 AA.
 ID GON3_PETMA
 AC P30948;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
 DE (Luliberin III).
 OS Petromyzon marinus (Sea lamprey).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 CC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=93178316; PubMed=8440174;
 RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
 RT "Primary structure and biological activity of a third gonadotropin-
 RT releasing hormone from lamprey brain.";
 RL Endocrinology 132:1125-1131(1993).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1277 MW; 284B36237AA1F5A3 CRC64;
 Query Match 19.4%; Score 31; DB 1; Length 10;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 23 HWSYGLRP 30
 Db 1
 |||: |

Db 2 HWSHWKP 9

RESULT 8

CT31_LITCI

ID CT31_LITCI STANDARD; PRT; 24 AA.
 AC P81851; P81852; P81853;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Citropin 3.1.2 [Contains: Citropin 3.1.1; Citropin 3.1].
 OS Litorea citropa (Australian blue mountains tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Litorea.
 OX NCBI_TaxID=94770;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin;
 RX MEDLINE=99435977; PubMed=10504394;
 RA Wegener K.L., Wabnitz P.A., Carver J.A., Bowie J.H., Chia B.C.S.,
 RA Wallace J.C., Tyler M.J.;
 RT "Host defence peptides from the skin glands of the Australian blue
 RT mountains tree-frog Litorea citropa. Solution structure of the
 RT antibacterial peptide citropin 1.1.";
 RL Eur. J. Biochem. 265:627-637(1999).
 CC -!- TISSUE SPECIFICITY: DORSAL AND SUBMENTAL SKIN GLANDS.
 FT PEPTIDE 1 24 CITROPIN 3.1.2.
 FT PEPTIDE 1 23 CITROPIN 3.1.1.
 FT PEPTIDE 1 22 CITROPIN 3.1.
 SQ SEQUENCE 24 AA; 2614 MW; C9001E295BD0E15D CRC64;

Query Match 19.4%; Score 31; DB 1; Length 24;
 Best Local Similarity 33.3%; Pred. No. 3e+02;
 Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGV 17
 | | : | : : | : | :
 Db 10 LKELTGGVIEGIGV 24

RESULT 9

COXJ_SHEEP

ID COXJ_SHEEP STANDARD; PRT; 24 AA.
 AC Q9TR30;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome c oxidase polypeptide VIIa-liver/heart, mitochondrial
 DE (EC 1.9.3.1) (Cytochrome c oxidase subunit VIIa-L), (Fragment).
 GN COX7A2 OR COX7AL.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver, and Heart;
 RX MEDLINE=96092035; PubMed=8529022;
 RA Linder D., Freund R., Kadenbach B.;
 RT "Species-specific expression of cytochrome c oxidase isozymes.";
 RL Comp. Biochem. Physiol. 112B:461-469(1995).
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
 KW Oxidoreductase; Inner membrane; Mitochondrion.
 FT NON_TER 24 24
 SQ SEQUENCE -24 AA; 2795 MW; D49D27C03B61F803 CRC64;

Query Match 18.4%; Score 29.5; DB 1; Length 24;
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;
 Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 KLLSEIKGVIVHRLGV 16
 | | : | : | : | : | : | :
 Db 10 KLFQEDNGIPVH-LKG 24

RESULT 10

ID PC24_BRANA STANDARD; PRT; 17 AA.
 AC P81097;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 20 kDa pollen coat protein (Fragment).
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. TOPAS; TISSUE=Pollen;
 RX MEDLINE=98345939; PubMed=9680961;
 RA Murphy D.J., Ross J.H.E.;
 RT "Biosynthesis, targeting and processing of oleosin-like proteins,
 RT which are major pollen coat components in Brassica napus.";
 RL Plant J. 13:1-16(1998).
 CC -!- FUNCTION: MAJOR COMPONENT OF THE POLLEN COAT.
 CC -!- TISSUE SPECIFICITY: POLLEN.
 CC -!- SIMILARITY: BELONGS TO THE OLEOSIN FAMILY.
 DR InterPro: IP8000136; Oleosin.
 DR PROSITE: PS00811; OLEOSINS; PARTIAL.
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA; 2018 MW; DC593833F02C52C9 CRC64;

Query Match 18.1%; Score 29; DB 1; Length 17;
 Best Local Similarity 42.9%; Pred. No. 4.1e+02;
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 LLSEIKGVIVHRL 15
 | | : | : | : | : | : | :
 Db 1 LQSPIRKIVNRK 14

RESULT 11

ID SODC_STRHE STANDARD; PRT; 31 AA.
 AC P81163;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) (Fragments).
 OS Striga hermonthica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Orobanchaceae; Striga.
 OX NCBI_TaxID=68872;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seedling;
 RA Stranger A., Corbett J.M., Dunn M.J., Totty N.F., Sterling A.,
 RA Bolwell G.P.;
 RT "Identification of developmentally-specific markers in germinating and
 RT haustorial stages of Striga hermonthica (Del.) Benth. seedlings.";
 RL J. Exp. Bot. 50:269-274(1999).
 CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
 CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -!- COFACTOR: Copper and zinc (By similarity).

```
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- DEVELOPMENTAL STAGE: PREFERENTIALLY EXPRESSED IN GERMINATING
CC SEEDLINGS.
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DR InterPro: IPR001424; SOD_CU_ZN.
DR PROSITE: PS00087; SOD_CU_ZN_1; PARTIAL.
DR PROSITE: PS00332; SOD_CU_ZN_2; PARTIAL.
KW Oxidoreductase; Copper; Zinc.
FT NON_TER 1
FT NON_CONS 17
FT NON_TER 31
FT NON_TER 31
SQ SEQUENCE 31 AA; 3109 MW; 40768DE8F8FC2958 CRC64;
Query Match 18.1%; Score 29; DB 1; Length 31;
Best Local Similarity 50.0%; Pred. No. 7.5e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 10 IVHRLGVEG 19
DB 4 VLHSEGVAG 13
RESULT 12
RS13_THETH
ID RS13_THETH STANDARD; PRT; 24 AA.
AC P80377;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 30S ribosomal protein S13 (Fragment).
GN RPSM OR RPS13
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE.
RX MEDLINE=95045586; PubMed=7957245;
RA Tsiiboli P., Herfurth E., Choli T.;
RT "Purification and characterization of the 30S ribosomal proteins from
RT the bacterium Thermus thermophilus.";
RL Eur. J. Biochem. 226:169-177(1994).
CC -!- FUNCTION: INVOLVED IN THE BINDING OF FMET-TRNA AND, HENCE, IN THE
CC INITIATION OF TRANSLATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro: IPR001892; Ribosomal_S13.
DR Pfam: PF00416; Ribosomal_S13; 1.
DR PROSITE: PS00646; RIBOSOMAL_S13; PARTIAL.
KW Ribosomal protein.
FT NON_TER 24
FT NON_TER 24
SQ SEQUENCE 24 AA; 2688 MW; 337F88EA0294D2A8 CRC64;
Query Match 17.8%; Score 28.5; DB 1; Length 24;
Best Local Similarity 34.8%; Pred. No. 6.8e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 7; Gaps 1;
QY 13 RLEGVEGP-----SLHWSYGL 28
DB 2 RIAGVEIPRKRVDVATYIIGI 24
RESULT 13
PETL_CYPAP
ID PETL_CYPAP STANDARD; PRT; 28 AA.
AC P48102;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome B6-F complex subunit VI (Cytochrome b6f complex subunit
DE petl).
GN PETL.
OS Cyanophora paradoxa.
OG Cyanelle.
CC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / PRINGSHEIM;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RA Bryant D.A.;
FT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RL Plant Mol. Biol. Rep. 13:327-332(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / PRINGSHEIM;
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RA Farley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,
RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RT the genetic complexity of a primitive plastid.";
RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL (1997).
CC -!- FUNCTION: THE CYTOCHROME B6-F COMPLEX FUNCTIONS IN THE LINEAR
CC CROSS-MEMBRANE TRANSPORT OF ELECTRONS BETWEEN PHOTOSYSTEM II AND
CC I, AS WELL AS IN CYCLIC ELECTRON FLOW AROUND PHOTOSYSTEM I. PETL
CC IS IMPORTANT FOR PHOTOAUTOTROPHIC GROWTH AS WELL AS FOR ELECTRON
CC TRANSFER EFFICIENCY AND STABILITY OF THE CYTOCHROME B6-F COMPLEX.
CC -!- SUBCELLULAR LOCATION: Thylakoid membrane-associated.
CC -!- SIMILARITY: BELONGS TO THE PETL FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: U30821; AAA81268.1;
KW Electron transport; Cyanelle; Respiratory chain; Transmembrane;
KW Thylakoid.
FT TRANSMEM 2 22
FT SEQUENCE 28 AA; 3106 MW; ADAE835D596AF3C CRC64;
Query Match 17.5%; Score 28; DB 1; Length 28;
Best Local Similarity 43.8%; Pred. No. 9.4e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 14 LEGVEGPSLHWSYGLR 29
DB 8 LSGMFGALACPFGLR 23
RESULT 14
AMAA_BACTR
ID AMAA_BACTR STANDARD; PRT; 24 AA.
AC P37356;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE N-acyl-L-amino acid amidohydrolase (EC 3.5.1.14) (L-aminoacylase)
DE (Fragment).
OS Bacillus thermoglucosidasius.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1426;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION
RA Cho H.-Y., Tanizawa K., Tanaka H., Soda K.;
RT "Thermotable aminoacylase from Bacillus thermoglucosidis.
RT Purification and characterization.";
RL Agric. Biol. Chem. 51:2793-2800(1987).
CC -!- CATALYTIC ACTIVITY: An N-acyl-L-amino acid + H(2)O = a fatty acid
CC anion + an L-amino acid.
CC -!- COFACTOR: ACTIVATED BY COBALT (BY SIMILARITY).
```

CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M40, ALSO KNOWN AS THE
 CC AMA/HIPO/HYUC FAMILY OF HYDROLASES.
 KW Hydrolase; Cobalt. 24
 FT NON_TER 24
 SQ SEQUENCE 24 AA; 2846 MW; BC954E4D2B0EC64A CRC64;

Query Match 17.2%; Score 27.5; DB 1; Length 24;
 Best Local Similarity 50.0%; Pred. No. 9.5e+02;
 Matches 7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 KLLSEIK-GVIVHR 13
 : : | : | : |
 Db 8 RLVDEVKEGVIAXR 21

RESULT 15
 COXK_SHEEP
 ID COXK_SHEEP STANDARD; PRT; 29 AA.
 AC Q9TR28;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome c oxidase polypeptide VIIa-heart, mitochondrial
 DE (EC 1.9.3.1) (Cytochrome c oxidase subunit VIIa-H) (COX VIIA-M)
 DE (Fragment).
 DE COX7A1 OR COX7AH.
 GN COX7A1 OR COX7AH.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Heart, and Liver;
 RX MEDLINE=96092035; PubMed=8529022;
 RA Linder D., Freund R., Kadenbach B.;
 RT "Species-specific expression of cytochrome c oxidase isozymes.";
 RL Comp. Biochem. Physiol. 112B:461-469(1995).
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
 KW Oxidoreductase; Inner membrane; Mitochondrion.
 FT NON_TER 29
 SQ SEQUENCE 29 AA; 3222 MW; 81712B0401B4373B CRC64;

Query Match 17.2%; Score 27.5; DB 1; Length 29;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 8; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Qy 1 KLLSEIKGVIVHRLEG 16
 : : | : | : |
 Db 10 KLFQEDNGLPVH-LKG 24

Search completed: October 10, 2002, 16:46:30
 Job time : 10.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: October 10, 2002, 16:44:12 ; Search time 19 Seconds
(without alignments)
282.255 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLGEGVSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 14138

Minimum DB seq length: 0
Maximum DB seq length: 31

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL19.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	23.8	27	12 Q9IIP8	Q9IIP8 hepatitis c
2	31	19.4	24	2 Q56150	Q56150 streptomyces
3	31	19.4	25	4 Q9UD88	Q9UD88 homo sapien
4	31	19.4	27	2 Q9P510	Q9P510 brevibacter
5	30.5	19.1	18	6 Q9PR15	Q9PR15 ovis aries
6	30.5	19.1	21	2 Q07840	Q07840 rhodobacter
7	30	18.8	18	4 Q9R4C3	Q9R4C3 agrobacteri
8	30	18.8	18	4 Q13767	Q13767 homo sapien
9	30	18.8	26	2 Q9ZAA4	Q9ZAA4 rhodobacter
10	30	18.8	31	4 Q96C35	Q96C35 homo sapien
11	29.5	18.4	26	2 Q30829	Q30829 campylobact
12	29	18.1	17	3 Q9H3E1	Q9H3E1 prochloroco
13	29	18.1	17	3 Q9H3E1	Q9H3E1 candida rug
14	29	18.1	27	12 Q9QI34	Q9QI34 hepatitis c
15	29	18.1	27	12 Q9QHW3	Q9QHW3 hepatitis c
16	29	18.1	27	12 Q9IIT6	Q9IIT6 hepatitis c

17	29	18.1	27	12 Q9IIT3	Q9IIT3 hepatitis c
18	29	18.1	27	12 Q9IIT1	Q9IIT1 hepatitis c
19	29	18.1	27	12 Q9IIS9	Q9IIS9 hepatitis c
20	29	18.1	27	12 Q9IIS6	Q9IIS6 hepatitis c
21	29	18.1	27	12 Q9IIS3	Q9IIS3 hepatitis c
22	29	18.1	30	4 Q9UBS6	Q9UBS6 homo sapien
23	29	18.1	30	4 Q96AP6	Q96AP6 homo sapien
24	29	18.1	31	8 Q9MS71	Q9MS71 leporocinlis
25	28	17.5	17	11 P97758	P97758 mus musculu
26	28	17.5	21	8 Q62950	Q62950 picea abies
27	28	17.5	25	4 Q9UQU9	Q9UQU9 homo sapien
28	28	17.5	27	8 Q94Q58	Q94Q58 schistosoma
29	28	17.5	27	12 Q9QIA9	Q9QIA9 hepatitis c
30	28	17.5	27	12 Q9QIA8	Q9QIA8 hepatitis c
31	28	17.5	27	12 Q9QIA7	Q9QIA7 hepatitis c
32	28	17.5	27	12 Q9QIA6	Q9QIA6 hepatitis c
33	28	17.5	27	12 Q9QIA5	Q9QIA5 hepatitis c
34	28	17.5	27	12 Q9QIA4	Q9QIA4 hepatitis c
35	28	17.5	27	12 Q9QIA2	Q9QIA2 hepatitis c
36	28	17.5	27	12 Q9QIA1	Q9QIA1 hepatitis c
37	28	17.5	27	12 Q9QIA0	Q9QIA0 hepatitis c
38	28	17.5	27	12 Q9QI99	Q9QI99 hepatitis c
39	28	17.5	27	12 Q9QI98	Q9QI98 hepatitis c
40	28	17.5	27	12 Q9QI97	Q9QI97 hepatitis c
41	28	17.5	27	12 Q9QI96	Q9QI96 hepatitis c
42	28	17.5	27	12 Q9QI94	Q9QI94 hepatitis c
43	28	17.5	27	12 Q9QI93	Q9QI93 hepatitis c
44	28	17.5	27	12 Q9QI91	Q9QI91 hepatitis c
45	28	17.5	28	4 Q9H4R8	Q9H4R8 homo sapien

ALIGNMENTS

RESULT 1					
Q9IIP8					
ID Q9IIP8	PRELIMINARY;	PRT;	27 AA.		
AC Q9IIP8:					
DT 01-OCT-2000 (TREMBlrel. 15, Created)					
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)					
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)					
DE ENVELOPE PROTEIN (FRAGMENT).					
OS Hepatitis C virus.					
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;					
OC Hepacivirus.					
OX NCBI_TaxID=111103;					
RN [1]					
RP SEQUENCE FROM N.A.					
RA Alberto S.-F.;					
RT "Influence of the dynamics of Hepatitis C virus quasiespecies in the					
RL histological outcome of liver transplantation.";					
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.					
DR EMBL; AF221397; AAF77959.1; -					
FT NON_TER 1					
FT NON_TER 27					
SQ SEQUENCE 27 AA; 2387 MW; 04A47C9B4D26C7C2 CRC64;					
Query Match 23.8%; Score 38; DB 12; Length 27;					
Best Local Similarity 39.1%; Pred. No. 1.4e+02;					
Matches 9; Conservative 1; Mismatches 9; Indels 4; Gaps 2;					
QY 8 GVIVHRLGEGVSLHWSYGLRP 30					
Db 6 GTAHTFAGLH-PSFWG---RP 24					
RESULT 2					
Q56150					
ID Q56150	PRELIMINARY;	PRT;	24 AA.		
AC Q56150;					
DT 01-NOV-1996 (TREMBlrel. 01, Created)					
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)					
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)					

DE HYPOTHETICAL 2.4 KDA PROTEIN (FRAGMENT).
OS Streptomyces violaceus (Streptomyces venezuelae).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1936;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISP5230;
RX MEDLINE=96070798; PubMed=7592948;
RA Mosher R.H., Camp D.J., Yang K., Brown M.P., Shaw W.V., Vining L.C.;
RT "Inactivation of chloramphenicol by O-phosphorylation. A novel
RT resistance mechanism in Streptomyces venezuelae ISP5230, a
RT chloramphenicol producer.";
RL J. Biol. Chem. 270:27000-27006(1995).
DR EMBL: U09991; AAB36571.1; -
KW Hypothetical protein..1
FT NON_TER 1
SQ SEQUENCE 24 AA; 2443 MW; 49DDCD3DB5FF0745 CRC64;
Query Match 19.4%; Score 31; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 19 GPSLHWSYG 27
II II I
Db 3 GPDPHWVC 11
RESULT 3
Q9UD88
ID Q9UD88 PRELIMINARY; PRT; 25 AA.
AC Q9UD88;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE APO(A) KRINGLE 4-37, APO(A) KRINGLE TYPE 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94060120; PubMed=7848387;
RA Pfaffinger D., McLean J., Scanu A.M.;
RT "Amplification of human APO(a) kringle 4-37 from blood lymphocyte
RT DNA.";
RL Biochim. Biophys. Acta 1225:107-109(1993).
DR HSP; P00747; ICRN.
SQ SEQUENCE 25 AA; 2928 MW; BEEDD4C62FA480A8 CRC64;
Query Match 19.4%; Score 31; DB 4; Length 25;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 20 PSLHWSY 26
II II I
Db 6 PSIRWEY 12
RESULT 4
Q9R510
ID Q9R510 PRELIMINARY; PRT; 27 AA.
AC Q9R510;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE NITRILE HYDRATASE BETA SUBUNIT (FRAGMENT).
OS Brevibacterium.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococineae; Brevibacteriaceae.
OX NCBI_TaxID=1696;
RN [1]
RP SEQUENCE. -

RX MEDLINE=92407758; PubMed=1527703;
RA Duran R., Chion C.K., Bigey F., Arnaud A., Galzy P.;
RT "The N-terminal amino acid sequences of Brevibacterium sp. R312
RT nitrile hydratase.";
RL J. Basic Microbiol. 32:13-19(1992).
DR HSP; P13449; 2AHJ.
SQ SEQUENCE 27 AA; 2743 MW; 17684C6389369073 CRC64;
Query Match 19.4%; Score 31; DB 2; Length 27;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 11 VHRLEGVEG 19
II II II I
Db 4 VHDLAGVQG 12
RESULT 5
Q9TRI5
ID Q9TRI5 PRELIMINARY; PRT; 18 AA.
AC Q9TRI5;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE SORBITOL DEHYDROGENASE (EC 1.1.1.14) (FRAGMENT).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RX MEDLINE=93170323; PubMed=8436142;
RA Reiersen H., Sletten K., McKinley-McKee J.S.;
RT "Affinity labelling of sorbitol dehydrogenase from sheep liver with
RT alpha-bromo-beta-(5-imidazolyl)propionic acid.";
RL Eur. J. Biochem. 211:861-869(1993).
SQ SEQUENCE 18 AA; 2076 MW; 10F3C812A111883A CRC64;
Query Match 19.1%; Score 30.5; DB 6; Length 18;
Best Local Similarity 38.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
QY 16 GVEGSLH-WSYG 27
II II II I
Db 5 GIXGSDVHWQHG 17
RESULT 6
O07840
ID O07840 PRELIMINARY; PRT; 21 AA.
AC O07840;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE UV ENDONUCLEASE (FRAGMENT).
GN UVRA.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2.4.1;
RX MEDLINE=98326818; PubMed=9663685;
RA Fernandez de Henestrosa A.R., Rivera E., Tapias A., Barbe J.;
RT "Identification of the Rhodobacter sphaeroides SOS box.";
RL Mol. Microbiol. 28:991-1003(1998).
DR EMBL: AF003108; AAC46417.1; -
KW Endonuclease.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2400 MW; B14ADAC2262DF490 CRC64;

Query Match¹ 19.18; Score 30.5; DB 2; Length 21;
Best Local Similarity 33.3%; Pred. No. 1.3e+03;
Matches 6; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Qy 1 KLLSEIKGVIVHLEGV 18
| : : : : | : : : :
Db 5 KFIS-VRGAREHNLKGD 21

RESULT 7

ID Q9R4C3 PRELIMINARY; PRT; 18 AA.
AC Q9R4C3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PROTOCATECHUATE 3,4-DIOXYGENASE TYPE I BETA SUBUNIT (EC 1.13.11.3) (FRAGMENT).
DE Agrobacterium tumefaciens.
OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE.
RX MEDLINE=96337865; PubMed=8772173;
RA Hammer A., Stolz A., Knackmuss H.;
RT "Purification and characterization of a novel type of protocatechuate
3,4-dioxygenase with the ability to oxidize 4-sulfocatechol.";
RL Arch. Microbiol. 166:92-100(1996).
SQ SEQUENCE 18 AA; 2008 MW; C7EBD971BBE5B8E9 CRC64;

Query Match 18.8%; Score 30; DB 2; Length 18;
Best Local Similarity 40.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 19 GPLSLHWSYGL 28
| : : | : :
Db 8 GPTFTWDXGI 17

RESULT 8

Q13767 Q13767 PRELIMINARY; PRT; 18 AA.
AC Q13767;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE ANTI-ANGIOTENSINOGEN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89273605; PubMed=2471521;
RA Moore G.J., Ganter R.C., Franklin K.J.;
RT "Angiotensin 'antipeptides': (-)messenger RNA complementary to human
angiotensin II (+)messenger RNA encodes an angiotensin receptor
antagonist.";
RL Biochem. Biophys. Res. Commun. 160:1387-1391(1989).
DR EMBL; M26228; AAA35530.1; -.
FT NON_TER 1
FT NON_TER 18
FT NON_TER 18
SQ SEQUENCE 18 AA; 1990 MW; A8E4F5A10C24F2CD CRC64;

Query Match 18.8%; Score 30; DB 4; Length 18;
Best Local Similarity 62.5%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 EIKGVIVH 12
| : : | : :
Db 6 EVEGVIVH 13

RESULT 9

Q9ZAA4 Q9ZAA4 PRELIMINARY; PRT; 26 AA.
AC Q9ZAA4;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE UVRA (FRAGMENT).
GN UVRA.
OS Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J50;
RX MEDLINE=99142327; PubMed=987839;
RA Labazi M., del Rey A., Fernandez de Henestrosa A.R., Barbe J.;
RT "A consensus sequence for the Rhodospirillaceae SOS operators.";
RL FEMS Microbiol. Lett. 171:37-42(1999).
DR EMBL; AF060210; AAC67504.1; -.
FT NON_TER 26
FT NON_TER 26
SQ SEQUENCE 26 AA; 2915 MW; C0263555EE10E534A CRC64;

Query Match 18.8%; Score 30; DB 2; Length 26;
Best Local Similarity 41.7%; Pred. No. 1.9e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 EIKGVIVHRL 16
| : : | : :
Db 8 EVRGAREHNLK 19

RESULT 10

Q96C35 Q96C35 PRELIMINARY; PRT; 31 AA.
AC Q96C35;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 3.9 KDA PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014871; AAHL4871.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 31
SQ SEQUENCE 31 AA; 3853 MW; BB8396CFAA7D0999 CRC64;

Query Match 18.8%; Score 30; DB 4; Length 31;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KLLSEIKGVIVHRL 14
| : : | : :
Db 11 QLLREIKGHKCYRI 24

RESULT 11

O30829 O30829 PRELIMINARY; PRT; 26 AA.
AC O30829;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ASPARTATE KINASE (FRAGMENT).

OS Campylobacter coli.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=195;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RMIT 32A;
 RA Linton D., Lawson A.J., Owen R.J., Stanley J.;
 RT "PCR detection, identification to species level, and fingerprinting of
 RT Campylobacter jejuni and Campylobacter coli direct from diarrheic
 RT samples";
 RL J. Clin. Microbiol. 35:0-0(1997).
 DR EMBL; AF017759; AAB82742.1; -;
 KW Kinase.
 FT NON_TER 1 1
 SQ SEQUENCE 26 AA; 3004 MW; 28EF62096D4C107F CRC64;

Query Match 18.1%; Score 29.5; DB 2; Length 26;
 Best Local Similarity 42.9%; Pred. No. 2.3e+03;
 Matches 9; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 9 VIVHLEGG-VEGSPSLHWSYGL 28
 Db 4 MIVHEKYGEIAVRAHCEYGL 24

RESULT 12

Q9X3E1 ID Q9X3E1 PRELIMINARY; PRT; 13 AA.

DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 GN CYTOCHROME B (FRAGMENT).
 GN PFTB
 OS Prochlorococcus sp.
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
 OC Prochlorococcus.
 OX NCBI_TaxID=1220;
 [1]
 RP SEQUENCE FROM N.A.
 RA Urbach E., Chisholm S.W.;
 RT "Genetic diversity in Prochlorococcus populations flow cytometrically
 RT sorted from the Sargasso Sea and Gulf Stream";
 RL Limnol. Oceanogr. 43:1615-1630(1998).
 DR EMBL; AF070141; AAD20755.1; -;
 FT NON_TER 1 1
 SQ SEQUENCE 13 AA; 1434 MW; 57EBE8029A8666D3 CRC64;

Query Match 18.1%; Score 29; DB 2; Length 13;
 Best Local Similarity 36.4%; Pred. No. 1.3e+03;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 IVHRLGEGP 20
 Db 2 LMSRKQGISGP 12

RESULT 13

Q9HDQ6 ID Q9HDQ6 PRELIMINARY; PRT; 17 AA.

DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE SPL1 TRNA SPLICING PROTEIN (FRAGMENT)
 OS Candida rugosa (Yeast) (Candida cylindracea).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5481;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATC3 14830;

RA Biasio W.;
 RL Thesis (2000), University of Vienna,
 RL Department of Biochemistry and Molecular Cell Biology.
 DR EMBL; AJ279020; CAC10273.1; -;
 FT NON_TER 1 1
 SQ SEQUENCE 17 AA; 1931 MW; C19ADBDB9B5455664 CRC64;
 Query Match 18.1%; Score 29; DB 3; Length 17;
 Best Local Similarity 36.4%; Pred. No. 1.7e+03;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 15 EGVEGSPSLHWS 25
 Db 5 EGIDLSTIWS 15

RESULT 14

Q9QI34 ID Q9QI34 PRELIMINARY; PRT; 27 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
 DE (NS1)] (FRAGMENT).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 [1]
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Izopet J.;
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 RT genome and the sensitivity of hepatitis C virus to interferon alpha
 RT therapy";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF166743; AAD52401.1; -;
 DR InterPro; IPR002531; HCV_NS1.
 DR Pfam; PF01560; HCV_NS1; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2833 MW; 218F8274D8FB7052 CRC64;

Query Match 18.1%; Score 29; DB 12; Length 27;
 Best Local Similarity 30.0%; Pred. No. 2.8e+03;
 Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 8 GVIVHRLGV 17
 Db 7 GIVAHNVKGL 16

RESULT 15

Q9QHY3 ID Q9QHY3 PRELIMINARY; PRT; 27 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
 DE (NS1)] (FRAGMENT).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 [1]
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Izopet J.;
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 RT genome and the sensitivity of hepatitis C virus to interferon alpha
 RT therapy";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF166794; AAD52452.1; -
DR InterPro; IPR002531; HCV_NS1.
DR Pfam; PF01560; HCV_NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2690 MW; 96F0689E426688A1 CRC64;

Query Match 18.1%; Score 29; DB 12; Length 27;
Best Local Similarity 30.8%; Pred. No. 2.8e+03;
Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 8 GVIVHRLEGVEGP 20
Db 7 GAVAHGARGITSP 19

Search completed: October 10, 2002, 16:48:27
Job time : 19 secs

PA (WANG/) WANG C Y.

PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX
 XX Claim 8; Page 84; 213pp; English.
 XX
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptens containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and haptens
 CC components. When the haptens is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX
 XX Sequence 25 AA;
 SQ
 Query Match 73.1%; Score 117; DB 15; Length 25;
 Best Local Similarity 85.7%; Pred. No. 8 6e-11;
 Matches 24; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 QY 3 LSEIKGVIVHRLEGVEGSLHWSYGLRP 30
 DB 1 LSEIKGVIVHRLEGVEGSLHWSYGLRP 24
 RESULT 2
 AAR62707
 ID AAR62707 standard; peptide; 27 AA.
 XX
 XX AAR62707;
 DT 10-SEP-1995 (first entry)
 XX
 XX LHRH-containing immunogenic peptide.
 XX
 XX Helper T cell epitope; universal immune stimulator; invasin; haptens;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Domain 1..15
 FT /note= "measles virus F protein helper T cell epitope"
 FT 18..27
 FT /note= "LHRH haptens"
 FT
 XX
 XX WC9425060-A.
 PN
 XX
 XX 10-NOV-1994.
 PD
 XX
 XX 28-APR-1994; 94WO-US04832.
 XX
 XX 27-APR-1993; 93US-0057166.
 PR
 XX 14-APR-1994; 94US-0229275.
 PR

XX (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX
 XX Claims 8, 12; Page 86; 213pp; English.
 XX
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptens containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and haptens
 CC components. When the haptens is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX
 XX This sequence is particularly preferred.
 XX
 XX Sequence 27 AA;
 SQ
 Query Match 72.5%; Score 116; DB 15; Length 27;
 Best Local Similarity 85.7%; Pred. No. 1.3e-10;
 Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
 QY 3 LSEIKGVIVHRLEGVEGSLHWSYGLRP 30
 DB 1 LSEIKGVIVHRLEGVEGSLHWSYGLRP 26
 RESULT 3
 AAY91156
 ID AAY91156 standard; peptide; 27 AA.
 XX
 XX AAY91156;
 DT 22-MAY-2000 (first entry)
 XX
 XX MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:36.
 DE
 XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.
 XX
 XX Chimeric - Measles virus.
 OS Chimeric - Rattus sp.
 OS
 XX WO9966957-A2.
 PN
 XX
 XX 29-DEC-1999.
 PD
 XX
 XX 21-JUN-1999; 99WO-US13975.
 PF
 XX


```

Db 1 LSEIKGVIVHRLGVGGE--HWSYGLRP 26
|||||
RESULT 5
AA91163
ID AA91163 standard; peptide; 27 AA.
XX
AC AA91163;
XX
DT 22-MAY-2000 (first entry)
XX
DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:43.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
KW cholesterol ester transport protein; anti-arteriosclerotic.
XX
OS Chimeric - Measles virus.
OS Chimeric - Rattus sp.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus
XX
PS Example 1; Page 80; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AA91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AA91122-Y91142,
CC AA91126 and AA911245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AA91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AA91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AA91156-Y91196, AA91127 and AA911242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AA91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AA91200 is somatostatin, and AA91201-Y91207 are antigenic

```

```

CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AA91208 is a
CC human CD4 CDR2-like domain antigenic site, and AA91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AA90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AA90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AA91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AA91221-Y91222 comprise this peptide and a Th
CC epitope. AA91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AA91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AA91228-Y91231 represent
CC CETP-derived peptides and AA91232-Y91241 are immunogens comprising a
CC CETP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AA91247 and AA91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AA91248-Y91251 and
CC AA91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AA91198 and AA91199 are respectively an immunostimulatory invasin
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.
XX
SQ Sequence 27 AA;
Query Match 70.6%; Score 113; DB 21; Length 27;
Best Local Similarity 82.1%; Pred. No. 3.8e-10;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
QY 3 LSEIKGVIVHRLGVGGEPSLHWSYGLRP 30
|||||
DB 1 LSEIKGVIVHRLGVGGE--HWSYGLRP 26
RESULT 6
AA91175
ID AA91175 standard; peptide; 31 AA.
XX
AC AA91175;
XX
DT 22-MAY-2000 (first entry)
XX
DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:55.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
KW cholesterol ester transport protein; anti-arteriosclerotic.
XX
OS Chimeric - Measles virus.
OS Chimeric - Rattus sp.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus
XX
PS Example 1; Page 80; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AA91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AA91122-Y91142,
CC AA91126 and AA911245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AA91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AA91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AA91156-Y91196, AA91127 and AA911242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AA91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AA91200 is somatostatin, and AA91201-Y91207 are antigenic

```


XX 29-DEC-1999.
XX 21-JUN-1999; 99WO-US13975.
XX 20-JUN-1998; 98US-0100412.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY;
XX WPI: 2000-160564/14.
XX New artificial T helper cell epitope and derived immunogens with target
XX antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX or human immune deficiency virus -
XX
XX Example 1; Page 86; 129pp; English.
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
XX and immunogenic peptides comprising the Th epitopes of the invention
XX along with B cell epitopes. The Th epitopes and peptide immunogens
XX containing them, are used to induce a T helper cell response,
XX specifically against Plasmodium falciparum, cholesterol ester transport
XX protein (CETP) or HIV epitopes, but more generally against any pathogen,
XX immunoreactive self-antigen or tumour antigen. The Th epitopes and
XX peptide immunogens may be used for prevention and/or treatment of
XX infections (HIV, foot-and-mouth disease or malaria); for cancer
XX immunotherapy; for inhibition of the action of luteinising hormone
XX releasing hormone (LHRH) for contraception, treatment of hormone-
XX dependent cancer, prevention of boar taint in meat, and
XX immunocastration); for promoting the growth of animals; or for
XX treating allergies or arteriosclerosis. Incorporation of a promiscuous
XX Th (functional in genetically diverse subjects) into an immunogen
XX improves capacity to induce a strong T helper cell-mediated immune
XX response, resulting in production of antibodies against a target
XX antigen. Th can replace carrier proteins and pathogen-derived T helper
XX epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
XX from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
XX AAY91126 and AAY91245-Y91246 represent synthetic Th epitopes based on the
XX MWF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
XX from hepatitis B virus (HBV) surface antigen, and sequences
XX AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
XX comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
XX is the LHRH target antigenic peptide used in these LHRH antigenic
XX peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
XX peptides comprising somatostatin and a Th epitope. Somatostatin
XX immunogens may be used to promote growth in livestock. AAY91208 is a
XX human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
XX epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
XX infection of T cells. AAY90212 is a modified version of a human IgE
XX (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
XX antigenic peptides which may be used in the treatment of allergies.
XX AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
XX VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
XX epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
XX antigen, and AAY91224-Y91225 comprise the CS antigen and an MWF Th
XX epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
XX CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a
XX CETP peptide and a Th epitope which may be used to prevent or treat
XX arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
XX are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
XX AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
XX epitope which may be used as a component in an anti-HIV-1 vaccine.
XX AAY91198 and AAY91199 are respectively an immunostimulatory invasive
XX protein epitope from *Yersinia* species, and hinge spacer peptide, both of
XX which may optionally be used in the antigenic peptides of the
XX invention.

SQ Sequence 31 AA;

Query Match 66.9%; Score 107; DB 21; Length 31;

Best Local Similarity 75.0%; Pred. No. 3.7e-09;
Matches 21; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 3 LSEIKGVIVHRLEGVGPDLHWSYGLRP 30
|||||||:|||||
Db 3 LSEIKGVIVHKLGMFLGGEHWSYGLRP 30
|||||||:|||||
RESULT 10
AAY91158
ID AAY91158 standard; peptide: 28 AA.
XX
XX AAY91158;
AC
XX
XX 22-MAY-2000 (first entry)
DT
XX
XX Modified MWF Th epitope/LHRH antigenic peptide, SEQ ID NO:38.
DE
XX
XX Promiscuous T-cell epitope: measles virus F protein; MWF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesterol ester transport protein; anti-arteriosclerotic.
XX
XX Chimeric - Measles virus.
OS
XX Chimeric - Rattus sp.
OS
XX WO9966957-A2.
PN
XX
XX 29-DEC-1999.
PD
XX
XX 21-JUN-1999; 99WO-US13975.
PF
XX
XX 20-JUN-1998; 98US-0100412.
PR
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
PA
XX
XX Wang CY;
PI
XX
XX WPI: 2000-160564/14.
DR
XX
XX New artificial T helper cell epitope and derived immunogens with target
XX antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX or human immune deficiency virus -
XX
XX Example 1; Page 78; 129pp; English.
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
XX and immunogenic peptides comprising the Th epitopes of the invention
XX along with B cell epitopes. The Th epitopes and peptide immunogens
XX containing them, are used to induce a T helper cell response,
XX specifically against Plasmodium falciparum, cholesterol ester transport
XX protein (CETP) or HIV epitopes, but more generally against any pathogen,
XX immunoreactive self-antigen or tumour antigen. The Th epitopes and
XX peptide immunogens may be used for prevention and/or treatment of
XX infections (HIV, foot-and-mouth disease or malaria); for cancer
XX immunotherapy; for inhibition of the action of luteinising hormone
XX releasing hormone (LHRH) for contraception, treatment of hormone-
XX dependent cancer, prevention of boar taint in meat, and
XX immunocastration); for promoting the growth of animals; or for
XX treating allergies or arteriosclerosis. Incorporation of a promiscuous
XX Th (functional in genetically diverse subjects) into an immunogen
XX improves capacity to induce a strong T helper cell-mediated immune
XX response, resulting in production of antibodies against a target
XX antigen. Th can replace carrier proteins and pathogen-derived T helper
XX epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
XX from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
XX AAY91126 and AAY91245-Y91246 represent synthetic Th epitopes based on the
XX MWF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
XX from hepatitis B virus (HBV) surface antigen, and sequences
XX AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
XX comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
XX is the LHRH target antigenic peptide used in these LHRH antigenic
XX peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
XX peptides comprising somatostatin and a Th epitope. Somatostatin
XX immunogens may be used to promote growth in livestock. AAY91208 is a
XX human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
XX epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
XX infection of T cells. AAY90212 is a modified version of a human IgE
XX (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
XX antigenic peptides which may be used in the treatment of allergies.
XX AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
XX VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
XX epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
XX antigen, and AAY91224-Y91225 comprise the CS antigen and an MWF Th
XX epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
XX CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a
XX CETP peptide and a Th epitope which may be used to prevent or treat
XX arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
XX are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
XX AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
XX epitope which may be used as a component in an anti-HIV-1 vaccine.
XX AAY91198 and AAY91199 are respectively an immunostimulatory invasive
XX protein epitope from *Yersinia* species, and hinge spacer peptide, both of
XX which may optionally be used in the antigenic peptides of the
XX invention.

CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 28 AA;

Query Match 66.2%; Score 106; DB 21; Length 28;

Best Local Similarity 64.3%; Pred. No. 4.6e-09;

Matches 18; Conservative 6; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGEGVSPSLHWSYGLRP 30

Db 2 ISEIKGVIVHRLGEGVSPSLHWSYGLRP 27

RESULT 11

ID AAY91173 standard; peptide; 31 AA.

XX AC AAY91173;

XX DT 22-MAY-2000 (first entry)

XX DE Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:53.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.

OS Chimeric - Rattus sp.

XX WO9966957-A2.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13975.

XX PR 20-JUN-1998; 98US-0100412.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX FI Wang CY;

XX DR WPI; 2000-160564/14.

XX

PT

PT New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus

PS Example 1; Page 83; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them are used to induce a T helper cell response.

CC specifically against Plasmodium falciparum, cholesterol ester, transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-

CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target

CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences

CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic

CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV

CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target

CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257

CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 31 AA;

Query Match 65.6%; Score 105; DB 21; Length 31;

Best Local Similarity 67.9%; Pred. No. 7.4e-09;

Matches 19; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGEGVSPSLHWSYGLRP 30

Db 3 ISEIKGVIVHRLGEGVSPSLHWSYGLRP 30

RESULT 12

AY68582

ID AAY68582 standard; peptide; 31 AA.

XX AC AAY68582;

XX 05-MAY-2000 (first entry)
XX Peptide Immunogen comprising a Th epitope and LHRH target antigen.
XX Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
XX Synthetic.
OS Measles virus.
OS Unidentified.
XX Key Location/Qualifiers
FH 1..19
FT Peptide /note= "helper Th epitope AAY68551"
FT Peptide 20..21
FT Peptide /note= "spacer"
FT Peptide 22..31
FT /note= "LHRH antigenic epitope AAY68566"
XX WO9966952-A1.
PN 29-DEC-1999.
PD 21-JUN-1999; 99WO-US13960.
XX 20-JUN-1998; 98US-0100414.
PR (UNBI-) UNITED BIOMEDICAL INC.
PA Wang CY;
PI WPI: 2000-160562/14.
DR New peptide immunogen containing luteinising hormone-releasing hormone
PT antigen site and helper T cell epitope, for e.g. contraception and
PT treatment of cancer -
XX Example 1; Page 79; 102pp; English.
XX The present sequence represents a peptide immunogen comprising a
CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
CC hormone-releasing hormone (LHRH). The synthetic Th epitope is derived
CC from a structured synthetic antigen library (SSAL) designated SSAL1 Th1.
CC SSAL Th1 is modelled after a promiscuous epitope taken from the F protein
CC of the Measles virus. The peptide immunogens cause induction of a
CC specific immune response to LHRH which is involved in regulation of a
CC spermatogenesis, ovulation, oestrus, sexual development and secretion
CC of sex hormones. Provision of a promiscuous T helper epitope (which is
CC functional in genetically diverse subjects) provides optimum
CC immunogenicity to the B cell epitopes of the target antigen and thus
CC high antibody titres against the target antigen. The peptide immunogens
CC of the invention are used to vaccinate against mammalian LHRH, for use
CC as (reversible) contraceptive; control of hormone-dependent tumours
CC (cancer of prostate or breast, also endometriosis); to prevent boar
CC taint (and improve meat quality) and for immunocastration.
XX Sequence 31 AA;
SQ
Query Match 65.6%; Score 105; DB 21; Length 31;
Best Local Similarity 67.9%; Pred. NO. 7.4e-09;
Matches 19; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 3 LSEIKGVIVHRLEGVEGSLHWSYGLRP 30
:|||||:|||||:|||||
DB 3 LSEIKGVIVHRLEGVEGSLHWSYGLRP 30
RESULT 13
AAY91170

ID AAY91170 standard; peptide; 27 AA.
XX AC AAY91170;
XX 22-MAY-2000 (first entry)
XX Modified MVF Th epitope/LHRH antigenic peptide, SEQ ID NO:50.
XX Promiscuous T-cell epitope; measles virus F protein; MVF;
DE hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
XX luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX Chimeric - Measles virus.
OS Chimeric - Rattus sp.
XX WO9966957-A2.
PN 29-DEC-1999.
PD 21-JUN-1999; 99WO-US13975.
XX 20-JUN-1998; 98US-0100412.
PR (UNBI-) UNITED BIOMEDICAL INC.
PA Wang CY;
PI WPI: 2000-160564/14.
DR New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX Example 1; Page 82; 129pp; English.
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response.
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone-
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, and AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3

CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MWF Th
 CC epitope, and AAY91226-Y91227 are immunogens comprising a
 CC CERP-derived peptide and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 27 AA;

Query Match 64.4%; Score 103; DB 21; Length 27;
 Best Local Similarity 67.9%; Pred. No. 1.3e-08;
 Matches 19; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
 Db 1 MSEMKGVIVHKGMEGGE--HWSYGLRP 26

RESULT 14

AAY68575
 ID AAY68575 standard; peptide; 27 AA.

XX AAY68575;

DT 05-MAY-2000 (first entry)

DE Peptide immunogen comprising a Th epitope and LHRH target antigen.

XX Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

XX Synthetic.
 OS Measles virus.
 OS Unidentified.

Key Location/Qualifiers
 FT Peptide 1..15 /note= "helper Th epitope AAY68546"
 FT Peptide 16..17 /note= "spacer"
 FT Peptide 17..27 /note= "LHRH antigenic epitope AAY68566"

XX WO9966952-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13960.

XX 20-JUN-1998; 98US-0100414.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160562/14.

XX New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer

XX Example 1; Page 72; 102pp; English.
 PS The present sequence represents a peptide immunogen comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The synthetic Th epitope is derived
 CC from a structured synthetic antigen library (SSAL) designated SSAL1 Th1.
 CC SSAL Th1 is modelled after a promiscuous epitope taken from the F protein
 CC of the Measles virus. The peptide immunogens cause induction of a
 CC specific immune response to LHRH which is involved in regulation of
 CC spermatogenesis, ovulation, oestrus, sexual development and secretion
 CC of sex hormones. Provision of a promiscuous T helper epitope (which is
 CC functional in genetically diverse subjects) provides optimum
 CC immunogenicity to the B cell epitopes of the target antigen and thus
 CC high antibody titres against the target antigen. The peptide immunogens
 CC of the invention are used to vaccinate against mammalian LHRH, for use
 CC as (reversible) contraceptive; control of hormone-dependent tumours
 CC (cancer of prostate or breast, also endometriosis); to prevent boar
 CC taint (and improve meat quality) and for immunocastration.

XX Sequence 27 AA;

Query Match 64.4%; Score 103; DB 21; Length 27;
 Best Local Similarity 67.9%; Pred. No. 1.3e-08;
 Matches 19; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
 Db 1 MSEMKGVIVHKGMEGGE--HWSYGLRP 26

RESULT 15

AAR62726
 ID AAR62726 standard; peptide; 28 AA.

XX AAR62726;

DT 17-SEP-1995 (first entry)

DE LHRH-containing immunogenic peptide.

XX Helper T cell epitope; universal immune stimulator; invasive; haptens;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW structured synthetic antigen library; SSAL.

XX Synthetic.

Key Location/Qualifiers
 FT Misc-difference 1 /label= Asp, Glu
 FT Misc-difference 2 /label= Leu, Ile, Val, Phe
 FT Misc-difference 4 /label= Glu, Asp
 FT Misc-difference 5 /label= Leu, Ile, Val, Phe
 FT Misc-difference 6 /label= Lys, Arg
 FT Misc-difference 8 /label= Leu, Ile, Val, Phe
 FT Misc-difference 9 /label= Leu, Ile, Val, Phe
 FT Misc-difference 10 /label= Leu, Ile, Val, Phe
 FT Misc-difference 12 /label= Lys, Arg
 FT Misc-difference 13 /label= Leu, Ile, Val, Phe
 FT Misc-difference 14 /label= Glu, Asp
 FT Misc-difference 16 /label= Leu, Ile, Val, Phe

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:47:38 ; Search time 11.5 Seconds
(without alignments)
65.843 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVIVHRLGVEGSPSLHWSYGLRPX 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 145535

Minimum DB seq length: 0
Maximum DB seq length: 31

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	72.5	27	1 US-08-446-692-19	Sequence 19, Appl
2	116	72.5	27	2 US-08-488-351A-19	Sequence 19, Appl
3	116	72.5	27	3 US-09-100-414B-36	Sequence 36, Appl
4	116	72.5	27	4 US-09-303-323-36	Sequence 36, Appl
5	113	70.6	27	5 US-09-100-414B-43	Sequence 43, Appl
6	113	70.6	27	6 US-09-303-323-43	Sequence 43, Appl
7	110	68.8	31	7 US-09-100-414B-55	Sequence 55, Appl
8	110	68.8	31	8 US-09-303-323-55	Sequence 55, Appl
9	108	67.5	27	9 US-09-100-414B-41	Sequence 41, Appl
10	108	67.5	27	10 US-09-100-414B-47	Sequence 47, Appl
11	108	67.5	27	11 US-09-303-323-41	Sequence 41, Appl
12	108	67.5	27	12 US-09-303-323-47	Sequence 47, Appl
13	107	66.9	31	13 US-09-100-414B-59	Sequence 59, Appl
14	107	66.9	31	14 US-09-303-323-59	Sequence 59, Appl
15	106	66.2	28	15 US-09-100-414B-38	Sequence 38, Appl
16	106	66.2	28	16 US-09-303-323-38	Sequence 38, Appl
17	105	65.6	31	17 US-09-100-414B-53	Sequence 53, Appl
18	105	65.6	31	18 US-09-303-323-53	Sequence 53, Appl
19	103.5	64.7	25	19 US-08-446-692-17	Sequence 17, Appl
20	103.5	64.7	25	20 US-08-488-351A-17	Sequence 17, Appl
21	103	64.4	27	21 US-09-100-414B-50	Sequence 50, Appl
22	103	64.4	27	22 US-09-303-323-50	Sequence 50, Appl
23	100	62.5	28	23 US-08-446-692-38	Sequence 38, Appl
24	100	62.5	28	24 US-08-488-351A-38	Sequence 38, Appl
25	99	61.9	28	25 US-09-100-414B-39	Sequence 39, Appl
26	99	61.9	28	26 US-09-303-323-39	Sequence 39, Appl
27	94	58.8	28	27 US-09-100-414B-37	Sequence 37, Appl

28 94 58.8 28 4 US-09-303-323-37 Sequence 37, Appl
29 84 52.5 27 3 US-09-100-414B-51 Sequence 51, Appl
30 84 52.5 27 4 US-09-303-323-51 Sequence 51, Appl
31 81 50.6 31 3 US-09-100-414B-64 Sequence 64, Appl
32 81 50.6 31 4 US-09-303-323-64 Sequence 64, Appl
33 79 49.4 27 3 US-09-100-414B-44 Sequence 44, Appl
34 79 49.4 27 3 US-09-100-414B-48 Sequence 48, Appl
35 79 49.4 27 3 US-09-100-414B-49 Sequence 49, Appl
36 79 49.4 27 4 US-09-303-323-44 Sequence 44, Appl
37 79 49.4 27 4 US-09-303-323-48 Sequence 48, Appl
38 79 49.4 27 4 US-09-303-323-49 Sequence 49, Appl
39 77 48.1 27 3 US-09-100-414B-42 Sequence 42, Appl
40 77 48.1 27 4 US-09-303-323-42 Sequence 42, Appl
41 76 47.5 28 3 US-09-100-414B-40 Sequence 40, Appl
42 76 47.5 28 4 US-09-303-323-40 Sequence 40, Appl
43 76 47.5 31 3 US-09-100-414B-56 Sequence 56, Appl
44 76 47.5 31 3 US-09-100-414B-61 Sequence 61, Appl
45 76 47.5 31 4 US-09-303-323-56 Sequence 56, Appl

ALIGNMENTS

RESULT 1
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-19

Query Match 72.5%; Score 116; DB 1; Length 27;
Best Local Similarity 85.7%; Pred. No. 2.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVEGSPSLHWSYGLRP 30
|||||
Db 1 LSEIKGVIVHRLGVEGSPSLHWSYGLRP 26

```
RESULT 2
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-19
Query Match 72.5%; Score 116; DB 2; Length 27;
Best Local Similarity 85.7%; Pred. No. 2.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
Oy 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26
|||||
RESULT 3
US-09-100-414B-36
; Sequence 36, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-36
Query Match 72.5%; Score 116; DB 3; Length 27;
Best Local Similarity 85.7%; Pred. No. 2.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
Oy 3 LSEIKGVIVHRLEGVEGSPSLHWSYGLRP 30
Db 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26
|||||
RESULT 4
US-09-303-323-36
; Sequence 36, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
```

```
;
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-36

Query Match 72.5%; Score 116; DB 4; Length 27;
Best Local Similarity 85.7%; Pred. No. 2.5e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVGEGPSLHWSYGLRP 30
   |||||
Db 1 LSEIKGVIVHRLGVGEGE--HWSYGLRP 26

RESULT 5
US-09-100-414B-43
; Sequence 43, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-43

Query Match 70.6%; Score 113; DB 3; Length 27;
Best Local Similarity 82.1%; Pred. No. 7.2e-11;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVGEGPSLHWSYGLRP 30
   |||||
Db 1 LSEIKGVIVHRLGVGEGE--HWSYGLRP 26

RESULT 6
US-09-303-323-43
; Sequence 43, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-43

Query Match 70.6%; Score 113; DB 3; Length 27;
Best Local Similarity 82.1%; Pred. No. 7.2e-11;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVGEGPSLHWSYGLRP 30
   |||||
Db 1 LSEIKGVIVHRLGVGEGE--HWSYGLRP 26

RESULT 7
US-09-100-414B-55
; Sequence 55, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-43

Query Match 70.6%; Score 113; DB 4; Length 27;
Best Local Similarity 82.1%; Pred. No. 7.2e-11;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVGEGPSLHWSYGLRP 30
   |||||
Db 1 LSEIKGVIVHRLGVGEGE--HWSYGLRP 26
```

NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-55

Query Match 68.8%; Score 110; DB 3; Length 31;
Best Local Similarity 78.6%; Pred. No. 2.4e-10;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
|||||
Db 3 LSEIKGVIVHKLVGFGGEHWSYGLRP 30

RESULT 8
US-09-303-323-55
Sequence 55, Application us/09303323
Patent No. 6228987

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 55:

SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-55

Query Match 68.8%; Score 110; DB 4; Length 31;
Best Local Similarity 78.6%; Pred. No. 2.4e-10;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
|||||
Db 3 LSEIKGVIVHKLVGFGGEHWSYGLRP 30

RESULT 9

US-09-100-414B-41
Sequence 41, Application US/09100414B
Patent No. 6025468

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-41

Query Match 67.5%; Score 108; DB 3; Length 27;
Best Local Similarity 71.4%; Pred. No. 4.2e-10;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLGVGSPSLHWSYGLRP 30
|||||
Db 1 LSEIKGVIVHKIEGIGE--HWSYGLRP 26

RESULT 10

US-09-100-414B-47
Sequence 47, Application US/09100414B
Patent No. 6025468

GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:


```
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-59

Query Match 66.9%; Score 107; DB 3; Length 31;
Best Local Similarity 75.0%; Pred. No. 7.le-10;
Matches 21; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLEGVPSLHWSYGLRP 30
Db 3 LSEIKGVIVHKLEGMFGGEHWSYGLRP 30

RESULT 14
US-09-303-323-59
; Sequence 59, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-758-4800
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-59

Query Match 66.2%; Score 106; DB 3; Length 28;
Best Local Similarity 64.3%; Pred. No. 8.9e-10;
Matches 18; Conservative 6; Mismatches 2; Indels 2; Gaps 1;

QY 3 LSEIKGVIVHRLEGVPSLHWSYGLRP 30
Db 2 ISEIRGIHRIEGIGGE--HWSYGLRP 27

Search completed: October 10, 2002, 16:51:26
Job time : 12 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:57:03 ; Search time 12.5 seconds
(without alignments)
261.363 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNNFTVSFWLRPKVSAHLEGPSLHWSYGLRXP 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 7455

Minimum DB seq length: 0

Maximum DB seq length: 34

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	28.0	10	1 RHPGG	gonadoliberin - pi
2	52	28.0	10	1 RHSHG	gonadoliberin - sh
3	48	25.8	10	1 RHAQ1	gonadoliberin I -
4	39	21.0	10	2 A21114	gonadoliberin - ch
5	36	19.4	30	2 S63531	hypothetical prote
6	34	18.3	10	1 RHAQ2	gonadoliberin II -
7	34	18.3	10	1 A61126	gonadoliberin - sp
8	34	18.3	10	2 A46030	gonadoliberin I -
9	34	18.3	10	2 B46030	gonadoliberin II -
10	32.5	17.5	26	2 T11816	hypothetical prote
11	32	17.2	32	2 A56589	galactose binding
12	31.5	16.9	29	2 I84189	cyclic AMP recepto
13	31.5	16.9	32	4 C26871	hypothetical chl p
14	31	16.7	10	2 A49187	gonadotropin-relea
15	31	16.7	16	2 S54271	GATA-2 protein - A
16	31	16.7	25	2 B36934	orf3 3' of mada -
17	31	16.7	29	2 B43937	endo-1,4-beta-xyla
18	31	16.7	30	2 S09518	prolamin - pearl m
19	30.5	16.4	30	2 S13753	replication initia
20	30.5	16.4	34	2 PC1319	large granule L9 c
21	30	16.1	25	2 I60083	glycophorin A - hu
22	30	16.1	34	2 D58493	group I allergen D
23	29.5	15.9	25	2 PH1716	Ig heavy chain v r
24	29.5	15.9	31	2 T01701	hypothetical prote
25	29.5	15.9	32	2 S20771	Ig heavy chain v r
26	29	15.6	17	2 I51203	myosin heavy chain
27	29	15.6	28	2 A36153	major allergen Ole
28	29	15.6	30	2 A49955	protein-tyrosine k
29	28.5	15.3	26	2 PQ0779	NADH dehydrogenase

T cell receptor be
lectin - Euphorbia
T-cell receptor be
hemoglobin beta x
Ig heavy chain v r
Ig heavy chain v r
Ig lambda chain v
hypothetical prote
toxic peptide pard
ribonuclease P1L -
microtubule-associ
probable rho leader
rho operon leader
rho operon leader
2S albumin small c
acetolactate synth

ALIGNMENTS

RESULT 1

RHPGG

gonadoliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Note: the synthetic and natural hormones have the same physicochemical and biologic

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyroglutamic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.0%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.27;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33

DB 2 HWSYGLRP 9

RESULT 2

RHSHG

gonadoliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

A;Residues: 1-10 <BUR>

A;Note: the natural and synthetic hormones have the same biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormones
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.0%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.27; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33

|||||

DB 2 HWSYGLRP 9

RESULT 3

RHAQ1

gonadoliberin I - American alligator
N;Alternate names: gonadotropin-releasing hormone I
C;Species: Alligator mississippiensis (American alligator)

C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997

C;Accession: A60066

R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson
Regul. Pept. 33, 105-116, 1991

A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of
A;Reference number: A60066; MUID:91352338

A;Accession: A60066

A;Molecule type: protein

A;Residues: 1-10 <LOV>

C;Superfamily: gonadoliberin

C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F;1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental

F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 25.8%; Score 48; DB 1; Length 10;

Best Local Similarity 87.5%; Pred. No. 0.98;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33

|||||

DB 2 HWSYGLRP 9

RESULT 4

A21114

gonadoliberin - chum salmon

C;Species: Oncorhynchus keta (chum salmon)

C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993

C;Accession: A21114

R;Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.

Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983

A;Title: Characterization of a teleost gonadotropin-releasing hormone.

A;Reference number: A21114; MUID:83195140

A;Accession: A21114

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <SHE>

Query Match 21.0%; Score 39; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33

|||||

DB 2 HWSYGLRP 9

RESULT 5

S63531

hypothetical protein 1 - Sulfolobus solfataricus (fragment)

C;Species: Sulfolobus solfataricus

C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999

C;Accession: S63531

R;Jones, C.E.; Fleming, T.M.; Cowan, D.A.; Littlechild, J.A.; Piper, P.W.

Eur. J. Biochem. 233, 800-808, 1995

A;Title: The phosphoglycerate kinase and glyceraldehyde-3-phosphate dehydrogenase genes

expression in Escherichia coli.

A;Reference number: S63528; MUID:96085144

A;Accession: S63531

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-30 <JON>

A;Cross-references: EMBL:X80178

Query Match 19.4%; Score 36; DB 2; Length 30;

Best Local Similarity 66.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 NFTVSFWLR 11

::: |||||

DB 5 SYFVSFWLR 13

RESULT 6

RHAQ2

gonadoliberin II - American alligator

N;Alternate names: gonadotropin-releasing hormone II

C;Species: Alligator mississippiensis (American alligator)

C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997

C;Accession: B60066

R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson

Regul. Pept. 33, 105-116, 1991

A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains

A;Reference number: A60066; MUID:91352338

A;Accession: B60066

A;Molecule type: protein

A;Residues: 1-10 <LOV>

C;Superfamily: gonadoliberin

C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F;1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental

F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 18.3%; Score 34; DB 1; Length 10;

Best Local Similarity 62.5%; Pred. No. 93;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33

|||||

DB 2 HWSHGWP 9

RESULT 7

A61126

gonadoliberin - spotted ratfish

N;Alternate names: gonadotropin-releasing hormone

C;Species: Hydroloagus collieri (spotted ratfish)

C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997

C;Accession: A61126

R;Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.-

Gen. Comp. Endocrinol. 82, 152-161, 1991

A;Title: Primary structure of gonadotropin-releasing hormone from the brain of a holbrookei

A;Reference number: A61126; MUID:91340067

A;Accession: A61126

A;Molecule type: protein

A;Residues: 1-10 <LOV>

A;Experimental source: brain

C;Superfamily: gonadoliberin

C;Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid

F;1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental

F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 18.3%; Score 34; DB 1; Length 10;

Best Local Similarity 62.5%; Pred. No. 93;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 26 HWSYGLRP 33
|||||
Db 2 HWSHGWP 9

RESULT 8

A46030
gonadoliberin I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 18.3%; Score 34; DB 2; Length 10;
Best Local Similarity 62.5%; Pred. No. 93;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 26 HWSYGLRP 33
|||||
Db 2 HWSHGWP 9

RESULT 9

B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 18.3%; Score 34; DB 2; Length 10;
Best Local Similarity 62.5%; Pred. No. 93;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 26 HWSYGLRP 33
|||||
Db 2 HWSHGWP 9

RESULT 10

Tl1816
hypothetical protein ORP26 - Norway spruce chloroplast
C:Species: Chloroplast Picea abies (Norway spruce)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: Tl1816
R:Kluemper, S.; Kanka, S.; Riesner, D.; Etscheid, M.
submitted to the EMBL Data Library, March 1997
A:Description: Characterisation of a Norway spruce chloroplast DNA clone: Complete nucle
A:Reference number: Z17349
A:Accession: Tl1816
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-26 <KLU>
A:Cross-references: EMBL:U92462; NID:g2959581; PID:g2959592
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 17.5%; Score 32.5; DB 2; Length 26;
Best Local Similarity 25.9%; Pred. No. 4.4e+02;
Matches 7; Conservative 6; Mismatches 7; Indels 7; Gaps 1;

Qy 7 SFWLRVPKVSASHLEGPSLHWSYGLRP 33
:|||||
Db 3 NFWIQSPAVIYP-----QPFPGVRP 22

RESULT 11

A56589
galactose binding lectin - beet armyworm (fragment)
C:Species: Spodoptera exigua (beet armyworm)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 13-Sep-1996
C:Accession: A56589
R:Boucias, D.G.; Pendland, J.C.
Insect Biochem. Mol. Biol. 23, 233-242, 1993
A:Title: The galactose binding lectin from the beet armyworm, Spodoptera exigua: disc
A:Reference number: A56589; MUID:93250866
A:Accession: A56589
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-32 <BOU>
A:Note: sequence extracted from NCBI backbone (NCBIP:131939)
C:Keywords: dimer; glycoprotein; hemolymph; lectin

Query Match 17.2%; Score 32; DB 2; Length 32;
Best Local Similarity 31.6%; Pred. No. 6.4e+02;
Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 3 NFTVSFWLRVPKVSASHLE 21
| : ||| : | : |
Db 7 NMDANGWLKVHQIPATWTE 25

RESULT 12

I84189
cyclic AMP receptor protein (CRP) - Escherichia coli
C:Species: Escherichia coli
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 08-Oct-1999
C:Accession: I84189
R:Kashiwagi, K.; Miyamoto, S.; Suzuki, F.; Kobayashi, H.; Igarashi, K.
Proc. Natl. Acad. Sci. U.S.A. 89, 4529-4533, 1992
A:Title: Excretion of putrescine by the putrescine-ornithine antiporter encoded by th
A:Reference number: I60729; MUID:92262473
A:Accession: I84189
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-29 <RES>
A:Cross-references: GB:M33766; NID:g806389; PIDN:AAA66175.1; PID:g455185

Query Match 16.9%; Score 31.5; DB 2; Length 29;
Best Local Similarity 44.4%; Pred. No. 6.8e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 5; Gaps 2;

Qy 9 WLRVPKVSASHLEGPSLH 26
||| : |||
Db 16 WLRI----SAHL-SPILH 28

RESULT 13

C26871
hypothetical chl protein - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: C26871
R:Johann, S.; Hinton, S.M.

J. Bacteriol. 169, 1911-1916, 1987
A:Title: Cloning and nucleotide sequence of the chld locus.
A:Reference number: A26871; MUID:87194564
A:Accession: C26871
A:Molecule type: DNA
A:Residues: 1-32 <JOH>
A:Cross-references: GB:M16182; NID:g145541; PID:AAA83841.1; PID:g1128946; GB:L34009; NI
A:Experimental source: strain K12
C:Comment: This is the hypothetical translation of a sequence that was not reported as a

Query Match 16.9%; Score 31.5; DB 4; Length 32;
Best Local Similarity 41.2%; Pred. No. 7.6e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 11 RVPKVSASHLEGPLHW 27
| :||| ||::|
Db 17 RCGRVSA---RGPGMNV 30

RESULT 14
A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A:Title: Primary structure and biological activity of a third gonadotropin-releasing hor
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A>Note: Sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 16.7%; Score 31; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33
|||: :|
Db 2 HWSHDWKP 9

RESULT 15
S54271
GATA-2 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995
C:Accession: S54271
R:Brewer, A.C.; Guille, M.J.; Fear, D.J.; Partington, G.A.; Patient, R.K.
EMBO J. 14, 757-766, 1995
A:Title: Nuclear translocation of a maternal CCAAT factor at the start of gastrulation a
A:Reference number: S54270; MUID:95188880
A:Accession: S54271
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-16 <BRE>

Query Match 16.7%; Score 31; DB 2; Length 16;
Best Local Similarity 58.3%; Pred. No. 4.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 11 RVPKVSASHLEG 22
| | | :|||
Db 4 RBPVSSSFHLEG 15

Search completed: October 10, 2002, 17:03:23
Job time : 13.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:52:48 ; Search time 9 Seconds
(without alignments)
146.274 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNNFTVSFWLRVPRKVSASHLEGPLSHWSYGLRPX 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 2222

Minimum DB seq length: 0

Maximum DB seq length: 34

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	48	25.8	10	GON1_ALLMI	P37041 alligator m
2	41	22.0	10	GON1_CLUPA	P81749 clupea gall
3	39	21.0	10	GON3_ONCKE	P20367 oncorhynch
4	34	18.3	10	GON2_CHICK	P37043 gallus gall
5	34	18.3	10	GONL_SQUAC	P27429 squalus aca
6	32	17.2	12	UR2_SCYCA	P35490 scyllorhinu
7	31	16.7	10	GON3_PETNA	P30948 petronyzon
8	30.5	16.4	29	GALA_SHEEP	P31234 ovis aries
9	28	15.1	18	AGI_EUPMA	P33889 euphorbia m
10	28	15.1	31	ER29_BOVIN	P81623 bos taurus
11	27.5	14.8	29	GALA_AMICA	P47214 amia calva
12	27	14.5	32	SODM_MYCHA	P80582 mycobacteri
13	27	14.5	32	ER29_CHICK	P81628 gallus gall
14	27	14.5	32	ER29_TRIVU	P81629 trichosurus
15	27	14.5	33	PAP1_PARPV	P81865 pardachirus
16	27	14.5	33	PAP4_PARMA	P81861 pardachirus
17	27	14.5	33	PAP5_PARMA	P81862 pardachirus
18	27	14.5	34	RNL1_PIG	P15466 sus scrofa
19	26.5	14.2	33	LPRH_ECOLI	P37324 escherichia
20	26	14.0	10	GON1_CHEPR	P80677 chelyosoma
21	26	14.0	14	FIBA_HORSE	P14452 equus cabal
22	26	14.0	23	IRBP_RABIT	P12664 eryctolagus
23	26	14.0	25	ACP_ERYLO	P80919 erythroba
24	26	14.0	26	PCW4_PACGO	P82426 pachycondyl
25	26	14.0	28	PW71_HCMWT	P24429 human cytom
26	26	14.0	33	PAP3_PARPV	P81866 pardachirus
27	25	13.4	8	ALL1_CYDPO	P82152 cydia pomon
28	25	13.4	10	GRP_RANRI	P23260 rana ridibu
29	25	13.4	23	GRP_ONCMY	O9ps30 oncorhynch
30	25	13.4	25	GRP_SCYCA	P09472 scyllorhinu
31	25	13.4	27	GRP_CANFA	P08989 canis famil
32	25	13.4	27	GRP_CHICK	P01295 gallus gall
33	25	13.4	27	GRP_PIG	P01294 sus scrofa

ALIGNMENTS

RESULT 1

GON1_ALLMI
ID GON1_ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberein I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I) (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC PIR; A60056; RHA01.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone, Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 25.8%; Score 48; DB 1; Length 10;

Best Local Similarity 87.5%; Pred. No. 0.8; Mismatches 1; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33

Db 2 HWSYGLQP 9

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

|||||:

```

OC Clupeinae; Clupea.
OX NCBI_TaxID=30724;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RL hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 AMIDATION.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;

Query Match 22.0%; Score 41; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 7; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33
DB 2 HWSHGLSP 9
|||||
|

RESULT 3
GN3_ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GnRH-III) (LH-
DE RH III) (Luliberin III).
GN GnRH3.
OS Oncorhynchus keta (Chum salmon), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O.keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RT hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR PIR: A21114; A21114.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

```

```

FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

Query Match 21.0%; Score 39; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33
DB 2 HWSYGLWP 9
|||||
|

RESULT 4
GN2_CHICK STANDARD; PRT; 10 AA.
ID GON2_CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GnRH-II)
DE (LH-RH II) (Luliberin II).
OS Gallus gallus (Chicken).
OS Alligator mississippiensis (American alligator),
OS Squatus acanthias (Spiny dogfish),
OS Hydrolagus colliet (Spotted ratfish) (Pacific ratfish), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=Chicken; TISSUE=Hypothalamus;
RX MEDLINE=84222059; PubMed=6427779;
RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
RA Matsuo H.;
RT "Identification of the second gonadotropin-releasing hormone in
RT chicken hypothalamus: evidence that gonadotropin secretion is
RT probably controlled by two distinct gonadotropin-releasing hormones
RT in avian species.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=A.mississippiensis; TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Luce V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
RT from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
RN [3]
RP SEQUENCE.
RC SPECIES=S.acanthias; TISSUE=Brain;
RX MEDLINE=92333500; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
RT dogfish brain provides insight into GnRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN [4]
RP SEQUENCE.
RC SPECIES=H.colliet; TISSUE=Brain;
RX MEDLINE=91340067; PubMed=1678723;
RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
RA Lee T.;
RT "Primary structure of gonadotropin-releasing hormone from the brain
RT of a holoccephalan (ratfish: Hydrolagus colliet).";
RL Gen. Comp. Endocrinol. 82:152-161(1991).
RN [5]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

```

```
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RT hormones, including a novel form, from an ancient teleost, herring.";
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR; B60066; RHAQ2.
DR PIR; A61126; A61126.
DR PIR; B46030; B46030.
DR InterPro: IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;

Query Match 18.3%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. NO. 62;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33
   |||:|:|
DB 2 HWSHGWP 9

RESULT 5
GNL_SQUAC
ID GNL_SQUAC STANDARD; PRT; 10 AA.
AC P27429;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
DE (Luliberin).
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalae; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92333500; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
RT dogfish brain provides insight into GNRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;

Query Match 18.3%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. NO. 62;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 26 HWSYGLRP 33
   |||:|:|
DB 2 HWSHGWP 9

RESULT 6
UR2_SCYCA
ID UR2_SCYCA STANDARD; PRT; 12 AA.
AC P35490;
```

```
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Urotensin II (U-II) (U-II).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RC TISSUE=Spinal cord;
RX MEDLINE=92319231; PubMed=1620290;
RA Conlon J.M., O'Harte F., Smith D.D., Balmert R.J., Hazon N.;
RT "Purification and characterization of urotensin II and parvalbumin
RT from an elasmobranch fish, Scyliorhinus canicula (common dogfish).";
RL Neuroendocrinology 55:230-235(1992).
CC -!- FUNCTION: HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC CORTICOTROPIN-RELEASING FACTOR. PROBABLY INVOLVED IN SMOOTH
CC MUSCLE STIMULATION.
CC -!- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.
DR InterPro: IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin_II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11
SQ SEQUENCE 12 AA; 1526 MW; 804729F9D579CEBA CRC64;

Query Match 17.2%; Score 32; DB 1; Length 12;
Best Local Similarity 62.5%; Pred. NO. 1.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 NNFTVSFW 9
   |||:|:|
DB 1 NNFSDCFV 8

RESULT 7
GN3_PETMA
ID GN3_PETMA STANDARD; PRT; 10 AA.
AC P30948;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
DE (Luliberin III).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
RT releasing hormone from lamprey brain.";
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237A1F5A3 CRC64;

Query Match 16.7%; Score 31; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. NO. 1.6e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 26 HWSYGLRP 33
DB 2 HWSHDWKP 9

RESULT 8
GALA_SHEEP
ID GALA_SHEEP STANDARD; PRT; 29 AA.
AC P31234;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Galanin.
GN GAL OR GALN OR GLNN.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92158824; PubMed=1724081;
RA Sillard R., Langel U., Joernvall H.;
RL Peptides 12:855-859(1991).
CC -!- FUNCTION: CONTRACTS SMOOTH MUSCLE OF THE GASTROINTESTINAL AND
CC GENTOURINARY TRACT, REGULATES GROWTH HORMONE RELEASE, MODULATES
CC INSULIN RELEASE, AND MAY BE INVOLVED IN THE CONTROL OF ADRENAL
CC SECRETION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GALANIN FAMILY.
DR InterPro; IPR001600; Galanin.
DR Pfam; PF01296; Galanin; 1.
DR PRINTS; PR00273; GALANIN.
DR ProDom; PD005962; Galanin; 1.
DR PROSITE; PS00861; GALANIN; 1.
KW Hormone; Neuropeptide; Amidation.
FT MOD_RES 29 29 AMIDATION.
SQ SEQUENCE 29 AA; 3185 MW; F718719B2D3FB089 CRC64;

Query Match 16.4%; Score 30.5; DB 1; Length 29;
Best Local Similarity 39.1%; Pred. No. 5.3e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 7; Gaps 1;

QY 16 SASHLEGP-----SLHWSYGL 31
DB 6 SAGYLLGPHADNHRSFHDKHGL 28

RESULT 9
AGI_EUPWA
ID AGI_EUPWA STANDARD; PRT; 18 AA.
AC P33889;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Galactose-inhibitable lectin (Fragment).
OS Euphorbia marginata.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Euphorbiaceae; Euphorbia.
OX NCBI_TaxID=28955;
RN [1]
RP SEQUENCE.
RC TISSUE=Latex;
RX MEDLINE=93357266; PubMed=8353129;
RA Stirpe F., Licastro F., Morini M.C., Parente A., Savino G.,
RA Abbondanza A., Bolognesi A., Falasca A.I., Rossi C.A.;
RT "Purification and partial characterization of a mitogenic lectin from
RT the latex of Euphorbia marginata.";
RL Biochim. Biophys. Acta 1158:33-39(1993).

-!- FUNCTION: LECTIN THAT BINDS GALACTOSE, GALACTOSE-CONTAINING SUGARS
AND GENIOBIOSE. IT IS STRONGLY MITOGENIC FOR HUMAN T LYMPHOCYTES
AND INDUCES THE RELEASE OF INTERLEUKIN-1 BETA AND TUMOR NECROSIS
FACTOR ALPHA FROM CULTURED MONONUCLEAR CELLS. IT HAS A STRONG
HEMAGGLUTININATING ACTIVITY.
-!- SUBUNIT: HOMODIMER.
-!- DEVELOPMENTAL STAGE: THE PRODUCTION OF THIS LECTING VARIES WITH
SEASONS BEING HIGHER IN LATE SPRING.
-!- PTM: N-GLYCOSYLATED.
-!- SIMILARITY: TO E.CHARACIAS LECTIN.
PIR: S36120; S36120.
KW Lectin; Glycoprotein.
FT VARIANT 1 1 MISSING (IN 20% OF THE CHAINS).
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1887 MW; 6F79F053FCC740AA CRC64;

Query Match 15.1%; Score 28; DB 1; Length 18;
Best Local Similarity 57.1%; Pred. No. 7.2e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 18 SHLEGPS 24
DB 5 SHISGPN 11

RESULT 10
ER29_BOVIN
ID ER29_BOVIN STANDARD; PRT; 31 AA.
AC P81623;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endoplasmic reticulum protein Erp29 (Fragments).
GN ERP29.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=98409275; PubMed=9738895;
RA Ferrari D.M., van Nguyen P., Kratzin H.D., Soeling H.D.;
RT "Erp28, a human endoplasmic-reticulum-luminal protein, is a member of
RT the protein disulfide isomerase family but lacks a CXXC thioredoxin-
RT box motif.";
RL Eur. J. Biochem. 255:570-579(1998).
CC -!- FUNCTION: DOES NOT SEEM TO BE A DISULFIDE ISOMERASE. PLAYS AN
CC IMPORTANT ROLE IN THE PROCESSING OF SECRETORY PROTEINS WITHIN THE
CC ER (BY SIMILARITY).
CC -!- SUBUNIT: DIMER.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
KW Endoplasmic reticulum.
FT NON_CONS 21 22
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 3454 MW; 30A976BC4A351E3B CRC64;

Query Match 15.1%; Score 28; DB 1; Length 31;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 TVSFWRVPK 14
DB 11 TVTFYKVIK 20

RESULT 11
GALA_AMICA
ID GALA_AMICA STANDARD; PRT; 29 AA.
AC P47214;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

```


DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Galanin.
 OS Amia calva (Bowfin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Amlifomes; Amlidae; Amia.
 OX NCBI_TaxID=7924;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Stomach; PubMed=7527531;
 RX MEDLINE=95083480;
 RA Wang Y., Conlon J.M.;
 RT "Purification and characterization of galanin from the
 RT phylogenetically ancient fish, the bowfin (*Amia calva*) and dogfish
 RT (*Scyliorhinus canicula*).";
 RL Peptides 15:981-986(1994).
 CC -!- FUNCTION: CONTRACTIS SMOOTH MUSCLE OF THE GASTROINTESTINAL AND
 CC GENITOURINARY TRACT, REGULATES GROWTH HORMONE RELEASE, MODULATES
 CC INSULIN RELEASE, AND MAY BE INVOLVED IN THE CONTROL OF ADRENAL
 CC SECRETION.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GALANIN FAMILY.
 CC InterPro: IPR001600; Galanin.
 DR Pfam: PF01296; Galanin; 1.
 DR PRINTS: PR00273; GALANIN.
 DR PRODOM: PD005962; Galanin; 1.
 DR PROSITE: PS00861; GALANIN; 1.
 KW Hormone; Neuropeptide; Amidation.
 FT MOD_RES 29 29
 FT SEQUENCE 29 AA; 3114 MW; 7518719B2D271627 CRC64;

 Query Match 14.8%; Score 27.5; DB 1; Length 29;
 Best Local Similarity 39.1%; Pred. No. 1.3e+03;
 Matches 9; Conservative 3; Mismatches 4; Indels 7; Gaps 1;

 Qy 16 SASHLEGP-----SLHWSYGL 31
 ||:||||| ||: ||
 Db 6 SAGYLGLPHAYDNRHSLNDRKHL 28

 RESULT 12
 SODM_MYCHA STANDARD; PRT; 18 AA.
 ID SODM_MYCHA
 AC R0582;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Mn] (EC 1.15.1.1) (Fragment).
 GN SODA OR SOD.
 OS Mycobacterium habana.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1784;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=TM-5135;
 RX MEDLINE=96262709; PubMed=8704977;
 RA Bisht D., Mehrotra J., Dhandra M.S., Singh N.B., Sinha S.;
 RT "A major T-cell-inducing cytosolic 23 kDa protein antigen of the
 RT vaccine candidate Mycobacterium habana is superoxide dismutase.";
 RL Microbiology 142:1375-1383(1996).
 CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
 CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -!- COFACTOR: Manganese (By similarity).
 CC -!- SUBUNIT: HOMODIMER (PROBABLE).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
 CC FAMILY.
 DR HSP: P17670; 1IDS.
 DR InterPro: IPR001189; SOD_M1.
 DR Pfam: PF00081; sodfe; 1.
 DR PROSITE: PS00088; SOD_MN; PARTIAL.
 KW Oxidoreductase; Manganese.

FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 1883 MW; A4161A3DAC93F710 CRC64;

 Query Match 14.5%; Score 27; DB 1; Length 18;
 Best Local Similarity 57.1%; Pred. No. 9.8e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

 Qy 23 PSLHWSY 29
 ||:|||||
 Db 6 PDLGWDY 12

 RESULT 13
 ER29_CHICK STANDARD; PRT; 32 AA.
 ID ER29_CHICK
 AC P81628;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Endoplasmic reticulum protein Erp29 (Fragment).
 GN ERP29.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RA Hubbard M.J.;
 RL Submitted (DEC-1998) to the SWISS-PROT data bank.
 CC -!- FUNCTION: DOES NOT SEEM TO BE A DISULFIDE ISOMERASE. PLAYS AN
 CC IMPORTANT ROLE IN THE PROCESSING OF SECRETORY PROTEINS WITHIN THE
 CC ER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
 KW Endoplasmic reticulum.
 FT NON_TER 32 32
 FT SEQUENCE 32 AA; 3764 MW; 41F96D8D9E831D7E CRC64;

 Query Match 14.5%; Score 27; DB 1; Length 32;
 Best Local Similarity 40.0%; Pred. No. 1.7e+03;
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

 Qy 5 TVSEFLRVPK 14
 ||:|||||
 Db 11 TITFVKVTPK 20

 RESULT 14
 ER29_TRIVU STANDARD; PRT; 32 AA.
 ID ER29_TRIVU
 AC P81629;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Endoplasmic reticulum protein Erp29 (Fragment).
 GN ERP29.
 OS Trichosurus vulpecula (Brush-tailed possum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
 OX NCBI_TaxID=9337;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RA Hubbard M.J.;
 RL Submitted (DEC-1998) to the SWISS-PROT data bank.
 CC -!- FUNCTION: DOES NOT SEEM TO BE A DISULFIDE ISOMERASE. PLAYS AN
 CC IMPORTANT ROLE IN THE PROCESSING OF SECRETORY PROTEINS WITHIN THE
 CC ER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
 KW Endoplasmic reticulum.
 FT NON_TER 32 32
 FT SEQUENCE 32 AA; 3764 MW; 41F96D8D9E831D7E CRC64;

Query Match 14.5%; Score 27; DB 1; Length 32;
Best Local Similarity 40.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 TVSFWLVRPK 14
|::|::|
Db 11 TITFYKVIK 20

RESULT 15

PAP1_PARPV
ID PAP1_PARPV STANDARD; PRT; 33 AA.
AC F81865;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pardaxin P-1.
OS Pardachirus pavoninus (Pacific sole) (Peacock sole).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Soleoidae; Soleidae; Pardachirus.
OX NCBI_TaxID=8286;
RN [1]
RP SEQUENCE.
RA Thompson S.A., Tachibana K., Nakanishi K., Kubota I.;
RT "Melittin-like peptides from the shark-repelling defense secretion of
RT the sole Pardachirus pavoninus.";
RL Science 233:341-343(1986).
CC -!- FUNCTION: EXHIBITS UNUSUAL SHARK REPELLENT AND SURFACTANT
CC PROPERTIES. FORMS VOLTAGE-DEPENDENT, ION-PERMEABLE CHANNELS IN
CC MEMBRANES. AT HIGH CONCENTRATION CAUSES CELL MEMBRANE LYSIS.
CC CAUSES DEATH IN KILLFISH ORYZIAS LATIPES IN 30 MINUTES AT A
CC CONCENTRATION OF 25 MICROGRAMS/ML.
CC -!- SUBUNIT: IN AQUEOUS SOLUTION EXISTS AS A TETRAMER.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: CONSISTS OF A C-TERMINAL HYDROPHILIC REGION AND A
CC PREDOMINANTLY HYDROPHOBIC REMAINDER.
CC -!- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.
KW Toxin.
SQ SEQUENCE 33 AA; 3396 MW; 5ABFFA0A16FC910D CRC64;

Query Match 14.5%; Score 27; DB 1; Length 33;
Best Local Similarity 38.5%; Pred. No. 1.8e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 8 FWLRVPKVSASHL 20
|::|::|
Db 2 FFALIPKIISPL 14

Search completed: October 10, 2002, 17:01:38
Job time : 10 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:54:03 ; Search time 19 Seconds
(without alignments)
309.570 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNNFTVSFWLKPVKVSASHLEGPLHWSYGLRPX 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 16130

Minimum DB seq length: 0

Maximum DB seq length: 34

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

1: sp-archaea:*

2: sp-bacteria:*

3: sp-fungi:*

4: sp-human:*

5: sp-invertebrate:*

6: sp-mammal:*

7: sp-mhc:*

8: sp-organelle:*

9: sp-phage:*

10: sp-plant:*

11: sp-rodent:*

12: sp-virus:*

13: sp-vertebrate:*

14: sp-unclassified:*

15: sp-rvirus:*

16: sp-bacteriap:*

17: sp-archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	22.0	33	13 Q9PT34	Q9pt34 oncorhynch
2	41	22.0	33	13 Q9W7G0	Q9w7g0 oncorhynch
3	39	21.0	33	12 Q91J04	Q91j04 tt virus. o
4	36	19.4	21	10 Q41566	Q41566 triticum ae
5	36	19.4	31	1 Q55314	Q55314 sulfolobus
6	35.5	19.1	29	6 Q9TRN1	Q9trn1 sus scrofa
7	34.5	18.5	33	12 Q69077	Q69077 human herpe
8	34	18.3	32	11 Q9CVW6	Q9cvw6 mus musculu
9	34	18.3	34	11 Q99KX7	Q99kx7 mus musculu
10	33	17.7	26	15 Q9JFP1	Q9jfp1 human t-cel
11	33	17.7	31	2 Q9KH08	Q9kh08 thermus aqu
12	33	17.7	33	12 Q91J13	Q91j13 tt virus. o
13	32.5	17.5	26	8 O62960	O62960 picea abies
14	32	17.2	13	2 Q9L8K1	Q9l8k1 enterococcu
15	32	17.2	22	11 Q9CVJ5	Q9cvj5 mus musculu
16	32	17.2	27	12 Q9IIP8	Q9iip8 hepatitis c

17	32	17.2	32	5 Q9TWU2	Q9twu2 spodoptera
18	32	17.2	33	6 Q951C4	Q951c4 sus scrofa
19	31.5	16.9	29	2 Q47650	Q47650 escherichia
20	31.5	16.9	32	12 Q9JG31	Q9jg31 tt virus. o
21	31.5	16.9	32	12 Q9JG27	Q9jg27 tt virus. o
22	31.5	16.9	33	12 Q91J10	Q91j10 tt virus. o
23	31	16.7	13	2 Q9WW72	Q9ww72 enterococcu
24	31	16.7	13	2 Q9WW71	Q9ww71 enterococcu
25	31	16.7	13	2 Q9X514	Q9x514 enterococcu
26	31	16.7	18	2 Q9LAP4	Q9lap4 enterococcu
27	31	16.7	24	2 Q56150	Q56150 streptomyce
28	31	16.7	25	4 Q9UD88	Q9ud88 homo sapien
29	31	16.7	26	15 Q9JFN8	Q9jfn8 human t-cel
30	31	16.7	29	2 Q9R5M6	Q9r5m6 streptomyce
31	31	16.7	33	12 Q91J17	Q91j17 tt virus. o
32	31	16.7	33	12 Q91J16	Q91j16 tt virus. o
33	31	16.7	33	12 Q91J08	Q91j08 tt virus. o
34	30	16.1	18	2 Q9R4C3	Q9r4c3 agrobacteri
35	30	16.1	18	4 Q96L04	Q96l04 homo sapien
36	30	16.1	19	10 Q40183	Q40183 lemna gibba
37	30	16.1	20	6 Q95MK6	Q95mk6 eulemur cor
38	30	16.1	20	6 Q95MK5	Q95mk5 varecia var
39	30	16.1	21	10 Q40181	Q40181 lemna gibba
40	30	16.1	25	4 Q9UM93	Q9um93 homo sapien
41	30	16.1	25	5 Q9NDS5	Q9nds5 drosophila
42	30	16.1	25	11 Q9QVA3	Q9qva3 rattus sp.
43	30	16.1	25	13 Q9PSC4	Q9psc4 xenopus lae
44	30	16.1	26	15 Q9JFQ3	Q9jfq3 human t-cel
45	30	16.1	26	15 Q9JFQ2	Q9jfq2 human t-cel

ALIGNMENTS

RESULT 1

Q9PT34	PRELIMINARY;	PRT;	33 AA.
ID	Q9PT34		
AC	Q9PT34;		
DC	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)		
DE	GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)		
DE	(LULIBERIN) (FRAGMENT)		
GN	GNRH1		
OS	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;		
OC	Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.		
OX	NCBI_TaxID=8022;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99312119; PubMed=10385393;		
RA	Von Schalburg K.R., Sherwood N.M.;		
RT	"Regulation and expression of gonadotropin-releasing hormone gene		
RT	differs in brain and gonads in rainbow trout.";		
RL	Endocrinology 140:3012-3024(1999).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Von Schalburg K.R., Sherwood N.M.;		
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.		
CC	-!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY		
CC	SIMILARITY).		
CC	-!- SIMILARITY: BELONGS TO THE GNRH FAMILY.		
DR	EMBL: AF110533; AADA3461.1; .		
DR	InterPro: IPR002047; AKH.		
DR	InterPro: IPR002012; GNRH.		
DR	Pfam: PF00446; GNRH; 1.		
DR	PROSITE: PS00256; AKH; UNKNOWN_1.		
DR	PROSITE: PS00473; GNRH; 1.		
KW	Amidation; Hormone.		
FT	NON_TER 33		
SQ	SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;		

Query Match 22.0%; Score 41; DB 13; Length 33;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 24 SLHWSYGLRP 33
Db 23 SQHWSYGWLP 32

RESULT 2

Q9W7G0 PRELIMINARY; PRT; 33 AA.
AC Q9W7G0;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
GN Oncochrychus mykiss (Rainbow trout) (Salmo gairdneri).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
RT differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110993; AAA43463.1; -.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;

Query Match 22.0%; Score 41; DB 13; Length 33;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 24 SLHWSYGLRP 33
Db 23 SQHWSYGWLP 32

RESULT 3

Q91J04 PRELIMINARY; PRT; 33 AA.
AC Q91J04;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE ORF2 (FRAGMENT).
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BABY-7;
RA Lin H.-H., Kao J.-H., Lee P.-I., Chen D.-S.;
RT "Early acquisition of TT virus in infants: Probable minor role of
RT maternal transmission.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF380386; AAK59300.1; -.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3712 MW; 36BD62745C5D3037 CRC64;

Query Match 21.0%; Score 39; DB 12; Length 33;
Best Local Similarity 37.0%; Pred. No. 1.2e+02;
Matches 10; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

OY 7 SFWLRVPKVSASHLEGPSSLHWSYGLRP 33
Db 3 SEWLSFPSPSAARTPRRGFHASGRVP 29

RESULT 4

Q41566 PRELIMINARY; PRT; 21 AA.
ID Q41566
AC Q41566;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HEAT SHOCK PROTEIN 16.9 (FRAGMENT).
GN HSP16.9-1LCL.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. MUSTANG;
RX MEDLINE=96382426; PubMed=8790290;
RA Joshi C.P., Nguyen H.T.;
RT "Differential display-mediated rapid identification of different
RT members of a multigene family HSP 16.9 in wheat.";
RL Plant Mol. Biol. 31:575-584(1996).
DR EMBL; L37071; AAA51391.1; -.
DR InterPro; IPR002068; Crystallin_HSP20.
DR Pfam; PF00011; HSP20; 1.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2248 MW; 522524E70B3320F4 CRC64;

Query Match 19.4%; Score 36; DB 10; Length 21;
Best Local Similarity 52.9%; Pred. No. 2e+02;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 12 VPKVSASHLEGPSSLHWS 28
Db 4 VPKVAKKPEVKSIHIS 20

RESULT 5

Q55314 PRELIMINARY; PRT; 31 AA.
ID Q55314
AC Q55314;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE URF2 PROTEIN (FRAGMENT).
GN URF2
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96085144; PubMed=8521845;
RA Jones C.E., Fleming T.M., Cowan D.A., Littlechild J.A., Piper P.W.;
RT "The phosphoglycerate kinase and glyceraldehyde-3-phosphate
RT dehydrogenase genes from the thermophilic archaeon Sulfolobus
RT solfataricus overlap by 8bp. Isolation, sequencing of the genes and
RT expression on Escherichia coli.";
RL Eur. J. Biochem. 233:800-808(1995).
RN [2]

DR	EMBL; M74710; AAA45789.1; -.
FT	NON_TER 1 1
FF	NON_TER 33 33
SQ	SEQUENCE 33 AA; 3812 MW; BBE97FD3321DAAE8 CRC64;
Query Match	18.5%; Score 34.5; DB 12; Length 33;
Best Local Similarity	40.08; Pred. No. 5.3e+02;
Matches	8; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
QY	11 RVPKVSASHLEGPS-LHWSY 29 II ::::I:: I : : I :
Dd	10 RVCELTASHVAPPVHGKY 29 II ::::I:: I : : I :
RESULT 8	
ID Q9CVM6	PRELIMINARY; PRT; 32 AA.
AC O9CVM6	
DT 01-JUN-2001	(TEMBUREl. 17, Created)
DD 01-JUN-2001	(TEMBUREl. 17, Last sequence update)
DE 01-JUN-2001	(TEMBUREl. 17, Last annotation update)
DE 181001OM0IRIK PROTEIN (FRAGMENT).	
Gn 181001OM0IRIK.	
OS Mus musculus (Mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OX Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBInfo_TaxID=10090;	
RN NC_010010.1 [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;	
RX Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Adachi J., Fukunishi Y., Konno H., Adachi J., Fukanaka S., Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukanaka S., Akizawa K., Izawa M., Nishi K., Kiyoisawa H., Kondo S., Yamanaoka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasakura T., Saito R., Sadato K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Glaser C., King B., Kochiya H., Kuehl P., Lewis S., Matsuo Y., Nikolaev I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Rasnick D., Akao T., Furuno M., Anon H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Humé D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeys P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.; Functional Annotation of a full-length mouse cDNA collection."	
RL Nature 409:685-690(2001).	
RR EMBL; AK007426; BAB25031.1; -.	
DR MGD; MG1.1916286; 181001OM0IRik.	
FT NON_TER 1	
SQ SEQUENCE 32 AA; 3511 MW; FE004DAAC90FA67A CRC64;	
Query Match	18.3%; Score 34; DB 11; Length 32;
Best Local Similarity	50.08; Pred. No. 6.1e+02;
Matches	6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY	16 SASHPGLPSLHW 27 I I : : I III
Dd	11 SGSAIDSLSLHW 22 I I : : I III
RESULT 9	
ID Q9KKX7	PRELIMINARY; PRT; 34 AA.
AC O9KKX7	
DT 01-JUN-2001	(TEMBUREl. 17, Created)
DD 01-JUN-2001	(TEMBUREl. 17, Last sequence update)
DE 01-DEC-2001	(TEMBUREl. 19, Last annotation update)
DE HYPOTHEtical 3.9 kDa protein.	

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY TUMOR. WAP-TGF ALPHA MODEL. 7 MONTHS OLD, GROSS
RC TISSUE;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: BC003965; RAH03965.1; -
KW Hypothetical protein.
SQ SEQUENCE 34 AA; 3942 MW; 91B5749F0184EDEB CRC64;

Query Match 18.3%; Score 34; DB 11; Length 34;
Best Local Similarity 30.6%; Pred. No. 6.5e+02;
Matches 11; Conservative 3; Mismatches 10; Indels 12; Gaps 2;

QY 6 VFWLRL-----VPKVSASHLEGPSTLHWSYGLR 32
: ||| | : ||| | : ||| | : ||| |
Db 1 MDLWLRQVLEEDDEEPPEGPPELEAPAL---YGAR 33

RESULT 10
Q9JFP1 PRELIMINARY; PRT; 26 AA.
ID Q9JFP1
AC Q9JFP1
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE TAX PROTEIN (FRAGMENT).
GN TAX.
OS Human T-cell lymphotropic virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2028;
RX MEDLINE=20036621; PubMed=10567638;
RA Talarmin A., Vion B., Ureta-Vidal A., Du Fou G., Marty C., Kazanji M.;
RT "First seroepidemiological study and phylogenetic characterization of
RT human T-cell lymphotropic virus type I and II infection among
RT Amerindians in French Guiana".
RL J. Gen. Virol. 80:3083-3088(1999).
DR EMBL: AF076263; AAF26832.1; -
FT NON_TER 1
SQ SEQUENCE 26 AA; 2965 MW; FD6E0EDC6A2DE83B CRC64;

Query Match 17.7%; Score 33; DB 15; Length 26;
Best Local Similarity 50.0%; Pred. No. 6.7e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 11 RVPKVSASHLEGPS 24
: ||| | : ||| | : ||| | : ||| |
Db 4 RQPSISGGLPPS 17

RESULT 11
Q9KH08 PRELIMINARY; PRT; 31 AA.
ID Q9KH08
AC Q9KH08
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROBABLE HISTIDYL TRNA SYNTHETASE (FRAGMENT).
GN HISRS.
OS Thermus aquaticus.
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX NCBI_TaxID=271;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20459059; PubMed=10862622;
RA Logan C.; Mayhew S.G.;

RT "Cloning, overexpression, and characterization of peroxiredoxin and
RT NADH peroxiredoxin reductase from *Thermus aquaticus*.";
RL J. Biol. Chem. 275:30019-30028(2000).
DR EMBL: AF276071; AAF82120.1; -
DR HSP; O32422; IQE0.
KW Aminoacyl-tRNA synthetase.
FT NON_TER 31
SQ SEQUENCE 31 AA; 3576 MW; E1B772791B961E0 CRC64;

Query Match 17.7%; Score 33; DB 2; Length 31;
Best Local Similarity 31.6%; Pred. No. 8.1e+02;
Matches 6; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 11 RVPKVSASHLEGPSTLHWSY 29
: ||| | : ||| | : ||| | : ||| |
Db 4 QIPRGTDVLPDSEKQWY 22

RESULT 12
Q91J13 PRELIMINARY; PRT; 33 AA.
ID Q91J13
AC Q91J13
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF2 (FRAGMENT).
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOTHER-3;
RA Lin H.-H., Kao J.-H., Lee P.-I., Chen D.-S.;
RT "Early acquisition of TT virus in infants: Probable minor role of
RT maternal transmission".
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF380377; AAK59291.1; -
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3769 MW; 36BD6268A2ACBA37 CRC64;

Query Match 17.7%; Score 33; DB 12; Length 33;
Best Local Similarity 33.3%; Pred. No. 8.7e+02;
Matches 9; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 7 SFWL RVPKVSASHLEGPSTLHWSYGLRP 33
: ||| | : ||| | : ||| | : ||| |
Db 3 SEWLSPRPSTARTPRREIRSRGRVP 29

RESULT 13
O62960 PRELIMINARY; PRT; 26 AA.
ID O62960
AC O62960;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF26.
OS Picea abies (Norway spruce) (Picea excelsa).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
OX NCBI_TaxID=3329;
RN [1]
RP SEQUENCE FROM N.A.
RA Klumpp S., Kanka S., Riesner D., Etscheid M.;
RT "Characterisation of a Norway spruce chloroplast DNA clone: complete
RT nucleotide sequences of rpl23, rpl2, rpl19, rpl22, rps3, trn I,
RT pseudo-ndhC and a residual inverted repeat B".
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL: U92462; AAC95506.1; -
KW Chloroplast.
SQ SEQUENCE 26 AA; 3023 MW; 47768D499EE8DC8C CRC64;

Query Match 17.5%; Score 32.5; DB 8; Length 26;
Best Local Similarity 25.9%; Pred. No. 7.9e+02;
Matches 7; Conservative 6; Mismatches 7; Indels 7; Gaps 1;
QY 7 SFWLRVPKVSASHLEGPLHWSYGLRP 33
:||||:| : :|:|
DB 3 NFWIOSPAVIYP-----QFPFGVRP 22

RESULT 14

Q9L8K1 PRELIMINARY; PRT; 13 AA.
AC Q9L8K1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE VANS (FRAGMENT).
GN VANSB.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM4281;
RX MEDLINE=20307504; PubMed=10846225;
RA Dahl K.H., Lundblad E.W., Roekenes T.P., Olsvik O., Sundsfjord A.;
RT "Genetic linkage of the vanB2 gene cluster to Tn5382 in vancomycin
RT resistant enterococci and characterization of two novel insertion
RT sequences.";
RL Microbiology 146:1469-1479(2000).
DR EMBL; AF201896; AAF73374.1; -.
FT NON_TER 1
SQ SEQUENCE 13 AA; 1502 MW; 8D0E282189F9672A CRC64;

Query Match 17.2%; Score 32; DB 2; Length 13;
Best Local Similarity 54.5%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 6 VSFWLRVPKVS 16
| ||| :| |
DB 1 VLFWLDLPPTS 11

RESULT 15

Q9CVJ5 PRELIMINARY; PRT; 22 AA.
AC Q9CVJ5;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE 1810049003RIK PROTEIN (FRAGMENT).
GN 1810049003RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boirelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007851; BAB25305.1; -.
DR MGD; MGI:1919517; 1810049003Rik.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2580 MW; 551140FDB3C83388 CRC64;
Query Match 17.2%; Score 32; DB 11; Length 22;
Best Local Similarity 45.5%; Pred. No. 7.7e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 7 SFWLRVPKVS 17
||| :| |
DB 5 SFWFKFPGTSS 15

Search completed: October 10, 2002, 17:02:37
Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:51:33 ; Search time 23 Seconds
(without alignments)
164.196 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVFWLRPKVKSASHLEGLSHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 294572

Minimum DB seq length: 0
Maximum DB seq length: 34

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	81.7	32	15 AAR62702	LHRH-containing im
2	118.5	63.7	31	21 AAY92655	PSMpep012 - P30 in
3	112	60.2	21	12 AAR11896	Immunogenic conjug
4	112	60.2	21	17 AAW06130	Tetanus toxoid pro
5	112	60.2	21	17 AAR88397	T-cell antigen TT3
6	112	60.2	21	18 AAW46449	Broad range helper
7	112	60.2	21	19 AAW67034	Tetanus toxin frag
8	112	60.2	21	20 AAW67579	T-cell epitope pep
9	112	60.2	21	20 AAW73222	Tetanus toxoid epi
10	112	60.2	21	21 AAB45512	Tetanus p30 epitop
11	112	60.2	21	21 AAY99876	Tetanus toxin T ce

12	112	60.2	21	21 AAY92626	Foreign epitope P3
13	112	60.2	21	21 AAY84428	Amino acid sequenc
14	112	60.2	21	21 AAY49260	CD4+ T cell epitop
15	112	60.2	21	22 AAE11764	Clostridium tetani
16	112	60.2	21	22 AAB85702	Amino acid sequenc
17	112	60.2	21	22 AAB85453	Universal tetanus
18	112	60.2	21	22 AAB61958	Tetanus Toxoid uni
19	112	60.2	21	22 AAB20144	Tetanus toxin T-ce
20	112	60.2	21	22 AAB68637	HR-2 B cell pepti
21	112	60.2	21	22 AAB46173	Tetanus toxoid TT9
22	112	60.2	21	22 AAB49072	Tetanus toxoid TT
23	112	60.2	22	22 AAB46176	Tetanus toxoid 947
24	112	60.2	31	21 AAY92653	PSMpep010 - P30 in
25	112	60.2	31	21 AAY92654	PSMpep011 - P30 in
26	112	60.2	33	22 AAB49075	Amyloid beta/tetan
27	111	59.7	34	16 AAR83562	IgE CH4 region con
28	108	58.1	32	16 AAR82597	IgE CH4 region con
29	107	57.5	22	15 AAR62693	Helper T cell epit
30	107	57.5	22	16 AAR82574	Tetanus toxin help
31	107	57.5	22	17 AAW05600	Tetanus toxin help
32	107	57.5	22	21 AAY80057	Pathogen derived T
33	107	57.5	22	21 AAY54540	T helper cell (Th)
34	107	57.5	22	21 AAY58769	Unidentified pepti
35	107	57.5	22	22 AAB84436	Amino acid sequenc
36	106	57.0	21	21 AAY96458	Tetanus toxin (TTD
37	103	55.4	19	21 AAY99061	HLA class II bindi
38	95.5	51.3	20	19 AAW69279	Tetanus toxin frag
39	94.5	50.8	20	19 AAW71322	Universal helper T
40	91	48.9	21	18 AAW35438	T-cell stimulatory
41	87.5	47.0	22	17 AAW03002	Carrier peptide fo
42	79	42.5	14	19 AAW78832	Tetanus toxoid pro
43	79	42.5	14	21 AAB33644	MHC class II assoc
44	79	42.5	14	22 AAG93740	Tetanus toxoid pep
45	79	42.5	15	19 AAW78833	Tetanus toxoid pro

ALIGNMENTS

RESULT 1	
AAR62702	ID AAR62702 standard; peptide; 32 AA.
XX	AC AAR62702;
XX	DT 10-SEP-1995 (first entry)
XX	DE LHRH-containing immunogenic peptide.
XX	OS Synthetic.
XX	Key Location/Qualifiers
FT Domain	1..22
FT Domain	/note= "tetanus toxin helper T cell epitope"
FT Domain	23..32
FT Domain	/note= "LHRH hapten"
XX	PN WO9425060-A.
XX	PD 10-NOV-1994.
XX	PF 28-APR-1994; 94WO-US04832.
XX	PR 27-APR-1993; 93US-0057166.
XX	PR 14-APR-1994; 94US-0229275.
XX	(LADD/) LADD A E.
PA	(WANG/) WANG C Y.

DR WPI; 1991-141874/20.
XX
XX Synthetic peptide(s) used as universal carriers - for preparing
PT immunogenic conjugates used as vaccines against Plasmodium
PT falciparum
XX
XX Claim 1; page 13; 16pp; English.
XX
XX This peptide corresponds to residues 947-967 of Tetanus toxin. It can
CC be used as a universal carrier for the prepn. of an immunogenic
CC conjugate. It is covalently bound to a peptide or polysaccharide
CC haptan derived from a pathogen. This conjugate can be used as
CC a vaccine for malaria. This peptide is recognised by different T-
CC helper cell clones in association with alleles of the human MHC.
CC It contains 2 epitopes: (a) 953-967, recognised by DR5-restricted
CC clones; and (b) 947-960, recognised by all other DR and DP-
CC restricted clones.
XX
XX
SQ Sequence 21 AA;
Query Match 60.2%; Score 112; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FNNFTVSFRLRPVKVSASHLE 21
Db 1 FNNFTVSFRLRPVKVSASHLE 21
RESULT 4
AAW06130
ID AAW06130 standard; Peptide; 21 AA.
XX
XX AAW06130;
AC
XX
DT 07-FEB-1997 (first entry)
XX
XX Tetanus toxoid protein T-cell epitope.
DE
XX
XX Cholesteryl ester transfer protein; CETP; antigen; vaccine;
KW cardiovascular disease; atherosclerosis; tetanus toxoid;
KW T-cell epitope.
XX
XX Clostridium tetani.
OS
XX
XX W09634888-A1.
PN
XX
PD 07-NOV-1996.
XX
XX 01-MAY-1996; 96WO-US06147.
PF
XX
PR 01-MAY-1995; 95US-0432483.
XX
XX (TCEL-) T CELL SCI INC.
PA
XX
XX Rittershaus CW, Thomas LJ;
PI
XX
XX WPI; 1996-506103/50.
DR
XX
XX Cholesteryl ester transfer protein B cell epitope linked to T cell
PT epitope - used to generate vaccine to regulate CETP activity for
PT decreasing the risk of developing a cardiovascular disease e.g.
PT atherosclerosis
XX
XX Claim 11; Page 43; 72pp; English.
PS
XX
XX A helper T-cell epitope (AAW06130) comprises amino acids 947-967
CC of tetanus toxoid protein. It can be utilised in novel peptide
CC vaccines (see also AAW06129, AAW06132) also including B-cell
CC epitope(s) from human or rabbit cholesteryl ester transfer
CC protein (CETP) to elicit an immune response against endogenous
CC CETP activity, thereby treating or preventing a cardiovascular
CC disease, such as atherosclerosis.

XX
SQ Sequence 21 AA;
Query Match 60.2%; Score 112; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FNNFTVSFRLRPVKVSASHLE 21
Db 1 FNNFTVSFRLRPVKVSASHLE 21
RESULT 5
AAR88397
ID AAR88397 standard; Peptide; 21 AA.
XX
XX AAR88397;
AC
XX
XX 12-JUN-1996 (first entry)
DT
XX
XX T-cell antigen TT3 peptide.
DE
XX
XX T-antigen; vaccine; antibody; T-cell; T-lymphocyte;
KW alpha-helix; coiled-coil heterodimer; core peptide; subunit.
KW
XX
XX Synthetic.
OS
XX
XX W09531480-A1.
PN
XX
XX 23-NOV-1995.
PD
XX
XX 18-MAY-1995; 95WO-CA00293.
PF
XX
XX 18-MAY-1994; 94US-0245507.
PR
XX
XX (SPIS-) SPI SYNTHETIC PEPTIDES INC.
PA
XX
XX Cachia PJ, Hodges RS, Houston ME, Irvin RT, Kay CW;
PI Zhou NE;
XX
XX WPI; 1996-010880/01.
DR
XX
XX Hetero:dimeric polypeptide immunogen in coiled-coil configuration
PT with different antigens on each sub:unit - useful in vaccines and
PT for antibody prodn.
XX
XX Claim 7; Page 62; 95pp; English.
PS
XX
XX This T-cell antigen TT3 peptide may be attached to a core peptide
CC contained in one of the 2 subunits of an alpha-helical coiled-coil
CC heterodimer. Each core peptide is comprised of terminal and
CC internal AA repeat sequences. This peptide antigen is attached
CC to the core peptide through covalent linkages to certain AA of the
CC internal repeats. The 2 subunits of the heterodimer are arranged
CC in a stable alpha-helical coiled-coil configuration having a 1:1
CC stoichiometry, and the peptide antigen is disposed toward the outer
CC surfaces of the configuration. The heterodimer may be used as a
CC synthetic vaccine (optionally multivalent) or to generate
CC antibodies.
XX
XX
SQ Sequence 21 AA;
Query Match 60.2%; Score 112; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 FNNFTVSFRLRPVKVSASHLE 21
Db 1 FNNFTVSFRLRPVKVSASHLE 21
RESULT 6
AAW46449

ID AAW46449 standard; Peptide; 21 AA.
AC AAW46449;
XX
DT 18-MAY-1998 (first entry)
XX
DE Broad range helper T cell epitope from the tetanus toxoid protein.
XX
XX Cholesteryl ester transfer protein; CETP; cholesteryl ester;
KW high density lipoprotein; HDL; very low density lipoprotein; VLDL;
KW low density lipoprotein; LDL; T cell epitope; antibody;
KW DNA plasmid-based vaccine; broad range helper T cell epitope;
KW treatment; cardiovascular disease.
XX
OS Clostridium tetani.
XX
XX WO9741227-A1.
PN
XX 06-NOV-1997.
PD
XX 01-MAY-1997; 97WO-US07294.
PF
XX 21-FEB-1997; 97US-0802967.
PR
XX 01-MAY-1996; 96US-0640713.
PR
XX (TCEL-) T CELL SCI INC.
PA
XX Thomas LJ;
PI
XX WPI; 1997-549731/50.
DR
XX DNA plasmid-based vaccine encodes CETP B cell and helper T cell
PT epitope(s) - used for elevating high density lipoprotein levels, and
PT for treating cardiovascular disease
PT
XX Disclosure; Page 44; 67pp; English.
XX
XX The present sequence represents a broad range helper T cell epitope
CC of the tetanus protein. It can be used in DNA plasmid-based vaccines
CC against cholesteryl ester transfer proteins (CETPs). CETPs mediate the
CC transfer of cholesteryl esters from high density lipoprotein (HDL)
CC to very low density lipoprotein (VLDL) and low density lipoprotein (LDL),
CC and vice versa. An increased CETP activity produces an atherogenic
CC lipoprotein profile and induces atherosclerosis. A DNA plasmid-based
CC vaccine comprises sequences encoding at least one B cell epitope of CETP
CC linked in frame with at least one segment encoding a broad range helper
CC T cell epitope. The vaccines can be used to elevate the ratio of
CC circulating HDL to circulating LDL, VLDL or total cholesterol in a human.
CC It can also be used for decreasing the level of endogenous CETP activity
CC in a human. The vaccine can be used to produce anti-CETP antibodies in
CC vivo and for treating cardiovascular disease.
XX
XX Sequence 21 AA;
SQ
Query Match 60.2%; Score 112; DB 18; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FNNFTVSEWLRVPKVSASHLE 21
DB 1 FNNFTVSEWLRVPKVSASHLE 21
RESULT 7
AAW67034
ID AAW67034 standard; peptide; 21 AA.
XX
AC AAW67034;
XX
DT 15-DEC-1998 (first entry)
XX
DE Tetanus toxin fragment (residues 947-967).
XX

KW Tetanus toxin; vaccine; antibody; carbohydrate peptide conjugate;
KW dendritic poly-lysine; epitope; tumour.
XX
OS Clostridium tetani.
XX
PN WO9843677-A1.
XX
PD 08-OCT-1998.
PD
XX 27-MAR-1998; 98WO-EP01922.
PF
XX 27-MAR-1997; 97US-0041726.
PR
XX (INSP) INST PASTEUR.
PA
XX Bay S, Cantacuzene D, Leclerc C, Lo-man R;
PI
XX WPI; 1998-557071/47.
DR
XX Carbohydrate peptide conjugate used as vaccine - comprises carrier
PT with dendritic poly-lysine enabling multiple epitopes to be
PT covalently attached
PT
XX Disclosure; Page 13; 55pp; English.
XX
XX The invention relates to a new carbohydrate peptide conjugate, which
CC comprises a carrier with a dendritic poly-lysine enabling multiple
CC epitopes to be covalently attached to it. Also claimed are: (1) an
CC antibody purified from biological fluid or cells of organisms
CC administered with the carbohydrate peptide conjugate, and (2) a diagnosis
CC kit comprising antigen-specific antibodies elicited by immunisation with
CC the carbohydrate peptide conjugate. The peptide conjugate, antibody and
CC diagnosis kit are used to provide pharmaceutical compositions and
CC vaccines against tumours. These can be used to support an immune response
CC against viral infections caused by hepatitis virus, HIV or cytomegalo
CC virus. They can be used to enhance immune responses, especially B- and T-
CC cell responses, of humans and animals against bacterial infections. The
CC carbohydrate peptide conjugate stimulates the antibody and T-cell
CC response without stimulating undesired immune responses. The composition
CC is capable of increasing the survival of tumour bearing humans and
CC animals. The present sequence corresponds to residues 947-967 of tetanus
CC toxin. The synthetic peptide corresponding to this sequence may be used
CC as an epitope in a carbohydrate peptide conjugate.
XX
XX Sequence 21 AA;
SQ
Query Match 60.2%; Score 112; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FNNFTVSEWLRVPKVSASHLE 21
DB 1 FNNFTVSEWLRVPKVSASHLE 21
RESULT 8
AAW67579
ID AAW67579 standard; peptide; 21 AA.
XX
AC AAW67579;
XX
DT 02-MAR-1999 (first entry)
XX
DE T-cell epitope peptide #5 for chimeric fimbria/T-cell epitope peptide.
XX
KW Chimeric; non-typable Haemophilus influenzae; fimbria; T-cell epitope;
KW immunogenic composition; immune response.
XX
OS Synthetic.
XX
XX US5843464-A.
PN
XX 01-DEC-1998.
PD

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 DB 1 FNNFTVSFWLRVPKVSASHLE 21
 RESULT 11
 AAY99876
 ID AAY99876 standard; Protein; 21 AA.
 XX AAY99876;
 AC AAY99876;
 XX 03-OCT-2000 (first entry)
 DT
 XX Tetanus toxin T cell epitope helper peptide P30.
 DE
 XX Human; MAGE-10; tumour rejection antigen precursor; bladder cancer;
 KW prostate cancer; lung cancer; cancer detection; oesophageal cancer;
 KW head and neck cancer; melanoma; myeloma; sarcoma; immunogen;
 KW tetanus toxin.
 XX Homo sapiens.
 OS
 XX WO200026407-A1.
 PN
 XX 11-MAY-2000.
 PD
 XX 15-OCT-1999; 99WO-US24258.
 PF
 XX 30-OCT-1998; 98US-0183714.
 PR
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Boon-Falleur T, Brasseur F, Rimoldi D, Deplaen E;
 PI
 XX WPI; 2000-451624/39.
 DR
 XX Determining presence of cancer in samples, especially useful for
 PT detecting bladder, prostate and lung cancer comprises assaying sample
 PT for expression of tumor rejection antigen precursor MAGE-10 -
 XX
 XX Example 12; Page 14; 26pp; English.
 PS
 XX The present sequence is a tetanus toxin T cell epitope known as
 CC Helper peptide P30. Hybrids of this peptide and an immunogenic peptide
 CC derived from tumour rejection antigen precursor MAGE-10 were used to
 CC generate polyclonal antiserum against MAGE-10. MAGE-10 binding monoclonal
 CC antibodies can be used to detect MAGE-10 expression. A correlation
 CC between MAGE-10 expression and cancer has been discovered and thus by
 CC determining the presence of MAGE-10, the presence of cancer can be
 CC determined. MAGE-10 expression can be detected using an immunoassay, an
 CC oligonucleotide hybridisation assay or via other standard techniques.
 CC This method is especially useful for determining the presence of
 CC bladder, oesophageal, head and neck, prostate or lung cancer, or
 CC melanoma, myeloma or sarcoma.
 XX
 XX Sequence 21 AA;

Query Match 60.2%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 12
 AAY92626
 ID AAY92626 standard; Protein; 21 AA.
 XX AAY92626;
 AC

XX 10-AUG-2000 (first entry)
 DT Foreign epitope P30.
 DE
 XX Foreign epitope P30.
 KW Foreign epitope; P2; prostate specific membrane antigen; PSM; Her2;
 KW Heregulin 2; Fibroblast growth factor 8b; FGF8b; vaccination;
 KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
 KW prostate cancer; cell-associated peptide antigen.
 XX Clostridium tetani.
 OS
 XX WO200020027-A2.
 PN
 XX 13-APR-2000.
 PD
 XX 05-OCT-1999; 99WO-DK00525.
 PF
 XX 05-OCT-1998; 98DK-0001261.
 PR
 XX 20-OCT-1998; 98US-0105011.
 PA (MEBI-) M & E BIOTECH AS.
 XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 PI
 XX WPI; 2000-349917/30.
 DR N-PSDB; AAA09461.
 DR
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 PT
 XX Example 1; Page 214; 220pp; English.
 PS
 XX The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, human
 CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
 CC simultaneous presentation by antigen producing cells (APCs) of the
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
 CC group derived from the PA and/or at least 1 B-cell group derived from the
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is
 CC foreign to the animal. Analogues of human PSM, human Her2 and
 CC human/murine FGF8b comprising a substantial part of all known and
 CC predicted CTL and B-cell epitopes of the respective PA and including at
 CC least one foreign T helper epitope (e.g. P2 and/or P30) are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer
 CC when the PA is human PSM, FGF8b and Her2, respectively.
 XX
 XX Sequence 21 AA;

Query Match 60.2%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 13
 AAY84428
 ID AAY84428 standard; Peptide; 21 AA.
 XX AAY84428;
 AC AAY84428;
 XX 25-JUL-2000 (first entry)
 DT
 XX Amino acid sequence of the tetanus toxoid P30 epitope.
 DE
 XX Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;
 KW tumour necrosis factor receptor; type II transmembrane protein;

KW osteoclast differentiation; CSF-1; osteoclast activator;
 KW immune response; osteoporosis; bone resorption;
 KW tetanus toxoid P30 epitope.

XX Clostridium tetani.

XX WO200015807-A1.

XX 23-MAR-2000.

XX 13-SEP-1999; 99WO-DK00481.

XX 15-SEP-1998; 98DK-0001164.

XX 02-OCT-1998; 98US-0102896.

XX (MEBI-) M & E BIOTECH AS.

XX Halkier T, Haaning J;

XX WPI; 2000-271444/23.

XX In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used
 PT to treat, prevent and ameliorate osteoporosis -

XX Example; Page 106; 110pp; English.

XX The present sequence represents the tetanus toxoid P30 epitope. It is
 CC used to create a fusion protein with murine osteoprotegerin ligand
 CC (OPGL). Osteoprotegerin is a secreted member of the tumour necrosis
 CC factor receptor family, which blocks osteoclastogenesis in a dose
 CC dependent manner. The OPGL protein is synthesised as a type II
 CC transmembrane protein. The murine and human OPGL polypeptides are 87%
 CC homologous. OPGL is a potent osteoclast differentiation factor when
 CC combined with CSF-1. It is not capable of inducing osteoclast
 CC differentiation in the absence of CSF-1. OPGL is also an activator of
 CC mature osteoclasts. The specification describes a method for the in vivo
 CC down-regulation of OPGL activity in an animal. The method comprises
 CC using at least one OPGL polypeptide or subsequence, and/or at least one
 CC OPGL analogue to induce an immune response in the animal. The method
 CC and OPGL polypeptide are useful for treating, preventing and ameliorating
 CC osteoporosis or other diseases or conditions characterised by excessive
 CC bone resorption.

XX Sequence 21 AA;

Query Match 60.2%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFRLRPVKVSASHLE 21
 Db 1 FNNFTVSFRLRPVKVSASHLE 21

RESULT 14

AA49260
 ID AAY49260 standard; peptide; 21 AA.

XX AAY49260;

XX 07-FEB-2000 (first entry)

XX CD4+ T cell epitope P30TT fragment.

XX Carrier protein; CD4+; T cell epitope; N6; N10; N19; immunogen;
 KW encapsulated bacteria.

XX Synthetic.

XX WO9955730-A2.

XX 04-NOV-1999.

PF 27-APR-1999; 99WO-IB00844.

XX 27-APR-1998; 98GB-0008932.

XX (CHIR-) CHIRON SPA.

XX Rappuoli R, Grandi G;

XX WPI; 2000-023325/02.

XX Carrier proteins containing CD4+ epitopes useful for protecting against
 PT diseases caused by encapsulated bacteria -
 XX Disclosure; Page 36; 76pp; English.

XX The invention provides carrier proteins comprising at least 5 CD4+ T
 CC cell epitope. The carrier protein comprises at least 1 of N6, N10 or
 CC N19. The carrier protein can be prepared by expressing a vector
 CC comprising a nucleic acid molecule encoding the protein, in a host cell
 CC and recovering the expressed protein. The carrier protein can also be
 CC produced by (a) constructing oligonucleotide molecules that encode
 CC peptide epitopes; (b) annealing the oligonucleotides to form duplexes;
 CC (c) introducing the duplexes into an expression vector; (d) introducing
 CC the expression vector into a host cell; and (e) isolating the fusion
 CC protein produced from a culture of the host cells. The carrier protein
 CC can be used as a protective immunogen in the control of diseases caused
 CC by encapsulated bacteria. Sequences AAY49256-266 represent CD4+ T cell
 CC epitopes inserted in the recombinant polypeptide carrier proteins.

XX Sequence 21 AA;

Query Match 60.2%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFRLRPVKVSASHLE 21
 Db 1 FNNFTVSFRLRPVKVSASHLE 21

RESULT 15

AAE11764
 ID AAE11764 standard; peptide; 21 AA.

XX AAE11764;

XX 18-DEC-2001 (first entry)

XX Clostridium tetani P30 epitope.

XX Amyloid protein; neuroprotective; nootropic; immunostimulant; vaccine;
 KW Alzheimer's disease; anticonvulsant; gene therapy; Pick's disease;
 KW antidiabetic; systemic amyloidosis; maturity onset diabetes; ALS;
 KW amyotrophic lateral sclerosis; Parkinson's disease; encephalopathy;
 KW Huntington's disease; fronto-temporal dementia; P30 epitope.

XX Clostridium tetani.

XX WO200162284-A2.

XX 30-AUG-2001.

XX 19-FEB-2001; 2001WO-DK00113.

XX 21-FEB-2000; 2000DK-0000265.

XX 01-MAR-2000; 2000US-186295P.

XX (MEBI-) M & E BIOTECH AS.

XX Birk P, Jensen MR, Nielsen KG;

XX WPI; 2001-589796/66.

XX N-PSDB; AAD18756.

xx In vivo down-regulation of amyloid protein for the treatment of
PT Alzheimer's, comprises presenting an amyloidogenic polypeptide or its
PT subsequence and/or at least one analogue of the amyloidogenic
PT polypeptide to the immune system -
xx
xx Example 3; Page 118; 120pp; English.
xx
CC The invention relates to a method for in vivo down-regulation of amyloid
CC protein such as beta amyloid (Abeta) in an animal, including human. The
CC method comprising presenting to the animal's immune system an
CC immunogenically effective amount of at least one amyloidogenic protein
CC or its subsequence and/or at least one analogue of the amyloidogenic
CC polypeptide. The amyloidogenic protein or its subsequence, and its
CC analogue is useful for the preparation of an immunogenic composition
CC comprising an adjuvant for down-regulating amyloid in an animal. They are
CC also useful in the treatment, prophylaxis or amelioration of Alzheimer's
CC disease or other diseases characterised by amyloid deposits. They are
CC also useful in the treatment of systemic amyloidosis, maturity onset
CC diabetes, Parkinson's disease, Huntington's disease, fronto-temporal
CC dementia, amyotrophic lateral sclerosis (ALS), Pick's disease and
CC prion-related transmissible spongiform encephalopathies. They are also
CC useful for inducing production of antibodies against an amyloidogenic
CC polypeptide. The present sequence is Clostridium tetani F30 epitope
CC related to the invention.
xx

SQ Sequence 21 AA;

Query Match 60.2%; Score 112; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

Search completed: October 10, 2002, 17:00:58
Job time : 23 secs


```
RESULT 2
US-08-488-351A-14
; Sequence 14, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Y1
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-14

Query Match 81.7%; Score 152; DB 2; Length 32;
Best Local Similarity 87.9%; Pred. No. 7.7e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 FNNFTVFWLRVPKVSASHLEGLSHSYGLRP 33
DB 3 FNNFTVFWLRVPKVSASHLEGLSHSYGLRP 31

RESULT 3
US-07-610-525-1
; Sequence 1, Application US/07610525
; Patent No. 5196512
; GENERAL INFORMATION:
; APPLICANT: BIANCHI Elisabetta
; APPLICANT: PESSI Antonello
; APPLICANT: CORRADIN Giampietto
; TITLE OF INVENTION: SYNTHETIC PEPTIDES USEFUL AS
```

```
; TITLE OF INVENTION: UNIVERSAL CARRIERS FOR THE PREPARATION OF IMMUNOGENIC CONJUGATES
; TITLE OF INVENTION: AND THEIR USE IN THE DEVELOPMENT OF SYNTHETIC VACCINES.
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHEA & GOULD
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10020-1193
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/610,525
; FILING DATE: 19901108
; CLASSIFICATION: 424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-827-3000
; TELEFAX: 212-840-6702
; TELEX: 423973
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acid residues
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; FRAGMENT TYPE: internal fragment
US-07-610-525-1

Query Match 60.2%; Score 112; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVFWLRVPKVSASHLE 21
DB 1 FNNFTVFWLRVPKVSASHLE 21

RESULT 4
US-08-661-052-12
; Sequence 12, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
```

ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MXI-043CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-661-052-12

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSEWLRVPKVSASHLE 21
|||||

Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 5

US-08-460-502-8
Sequence 8, Application US/08460502
Patent No. 5843464

GENERAL INFORMATION:
APPLICANT: Bakaletz, Lauren O.
APPLICANT: Kaumaya, Parvin T.
TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter and Griswold
STREET: 800 Superior Avenue
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2688

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Goldrick, Mary E.
REGISTRATION NUMBER: 34,829
REFERENCE/DOCKET NUMBER: 22727/00120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8458
TELEFAX: (216) 241-0816
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-502-8

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSEWLRVPKVSASHLE 21
|||||

Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 6
US-08-724-774B-5
Sequence 5, Application US/08724774B
Patent No. 5908778
GENERAL INFORMATION:
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie,
APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
APPLICANT: Stefan; Reed, Daryl
TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
TITLE OF INVENTION: Rejection Antigen Precuros MAGE-10,
TITLE OF INVENTION: Antibodies Specific To The Molecule, and
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,774B
FILING DATE: 03-October-1996
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5908778man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: amino acid
TOPOLOGY: linear
US-08-724-774B-5

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSEWLRVPKVSASHLE 21
|||||

Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 7

US-09-089-595-5
Sequence 5, Application US/09089595
Patent No. 6153728

GENERAL INFORMATION:
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie,
APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
APPLICANT: Stefan; Reed, Daryl
TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
TITLE OF INVENTION: Rejection Antigen Precuros MAGE-10,
TITLE OF INVENTION: Antibodies Specific To The Molecule, and
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA

ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/089,595
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,774
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6153728man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: amino acid
TOPOLOGY: linear
US-09-089-595-5

Query Match 60.2%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 8

US-09-382-855-5
Sequence 5, Application US/09382855
Patent No. 6174692
GENERAL INFORMATION:
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie, Pierre;
APPLICANT: Carrozzini, Jean-Charles; Carrel, Stefan; Reed, Daryl
TITLE OF INVENTION: MAGE-10 ENCODING CDNA, The Tumor Rejection
TITLE OF INVENTION: Antigen Precursors Mage-10, Antibodies Specific To The Molecule,
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/382,855
FILING DATE: 25-August-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/089,595
FILING DATE: 02-June-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,774
FILING DATE: 03-October-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Schofield, Mary Anne
REGISTRATION NUMBER: 36,669
REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/WAS
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: amino acid
TOPOLOGY: linear
US-09-382-855-5

Query Match 60.2%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 9

US-09-183-714B-5
Sequence 5, Application US/09183714B
Patent No. 6221593
GENERAL INFORMATION:
APPLICANT: Boon-Falleur, Thierry
APPLICANT: Brasseur, Francis
APPLICANT: Rimoldi, Donata
APPLICANT: De Plaen, Etienne
TITLE OF INVENTION: Method for Determining Cancer by Determining Expression
TITLE OF INVENTION: of MAGE-10
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/183,714B
CURRENT FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: US 08/724,774
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 7
SEQ ID NO 5
LENGTH: 21
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
US-09-183-714B-5

Query Match 60.2%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10

US-09-188-082-12
Sequence 12, Application US/09188082
Patent No. 6270765
GENERAL INFORMATION:
APPLICANT: Yashwant M. Deo
APPLICANT: Joel Goldstein
APPLICANT: Robert Graziano
APPLICANT: Chezian Somasundaram
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/188.082
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/661.052
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MXI-043CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-09-188-082-12

Query Match 60.2%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSEWLRVPKVSASHLE 21
|||||
Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 11
US-09-171-969-10
Sequence 10, Application US/09171969
Patent No. 6284533
GENERAL INFORMATION:
APPLICANT: Thomas, Lawrence J.
TITLE OF INVENTION: PLASMID-BASED VACCINE FOR TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 75 State Street, Suite 2300
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1807
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171.969
FILING DATE: 01 May 1997 (01.05.97)
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/640.713
FILING DATE: 01 May 1996 (01.05.96)
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/802.967
FILING DATE: 21 February 1997 (21.02.97)
ATTORNEY/AGENT INFORMATION:
NAME: Leon R. Yankwich
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

HYPOTHETICAL:
ANTI-SENSE:
FEATURE:
NAME/KEY:
LOCATION:
US-09-171-969-10

Query Match 60.2%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSEWLRVPKVSASHLE 21
|||||
Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 12
PCT-US93-11703-66
Sequence 66, Application PC/TUS9311703
GENERAL INFORMATION:
APPLICANT: Chiron Mimotopes Pty. Ltd.
TITLE OF INVENTION: T-Cell Epitopes
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grant D. Green
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11703
FILING DATE: 28-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,852
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0222.101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-11703-66

Query Match 60.2%; Score 112; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSEWLRVPKVSASHLE 21
|||||
Db 1 FNNFTVSEWLRVPKVSASHLE 21

RESULT 13
PCT-US93-11703-64
Sequence 64, Application PC/TUS9311703
GENERAL INFORMATION:
APPLICANT: Chiron Mimotopes Pty. Ltd.
TITLE OF INVENTION: T-Cell Epitopes

NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grant D. Green
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11703
FILING DATE: 28-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,852
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0222.101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-11703-64

Query Match 60.2%; Score 112; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
|||||

DB 7 FNNFTVSFWLRVPKVSASHLE 27

RESULT 14
US-08-446-692-5
Sequence 5, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446.692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-5

Query Match 57.5%; Score 107; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHL 20
|||||

DB 3 FNNFTVSFWLRVPKVSASHL 22

RESULT 15
US-08-488-351A-5
Sequence 5, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488.351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446.692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229.275
FILING DATE: 14-APR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057.166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-5
Query Match      57.5%; Score 107; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 FNNFTVSFWLRVPKVSASHL 20
    |||||
Db   3 FNNFTVSFWLRVPKVSASHL 22
    |||||

Search completed: October 10, 2002, 17:04:05
Job time : 11 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:05:27 ; Search time 16 Seconds
(without alignments)
168.156 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKFIGITELGSLHWSYGLRDX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 5685

Minimum DB seq length: 0

Maximum DB seq length: 28

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	35.1	10	1 RHPGG	gonadoliberin - pi
2	52	35.1	10	1 RSHG	gonadoliberin - sh
3	48	32.4	10	1 RHAQ1	gonadoliberin I -
4	39	26.4	10	2 A2114	gonadoliberin - ch
5	35	23.6	27	2 T01664	envelope protein -
6	34	23.0	10	1 RHAQ2	gonadoliberin II -
7	34	23.0	10	1 A61126	gonadoliberin - sp
8	34	23.0	10	2 A46030	gonadoliberin I -
9	34	23.0	10	2 A46030	gonadoliberin II -
10	31	20.9	10	2 A49187	gonadotropin-relea
11	31	20.9	22	2 PQ0070	T-cell receptor be
12	29	19.6	23	2 PC4030	rRNA endonuclease
13	28.5	19.3	21	2 S28436	major outer membra
14	28	18.9	16	2 S57517	T cell receptor be
15	28	18.9	20	2 S57286	translation elonga
16	28	18.9	21	2 JU0386	nitrite hydratase
17	28	18.9	24	2 PT0258	Ig heavy chain CDR
18	27.5	18.6	15	2 P50455	superoxide dismuta
19	27	18.2	25	2 B36934	orf3 3' of mada -
20	26.5	17.9	27	1 RUCHA	gastrin-releasing
21	26.5	17.9	27	2 C85939	hypothetical prote
22	26	17.6	17	2 S58660	H+-transporting AT
23	26	17.6	22	2 PH1325	Ig heavy chain DJ
24	26	17.6	23	2 S51188	aldehyde dehydroge
25	26	17.6	25	2 A50286	heat-stable serine
26	26	17.6	26	2 S32248	Ig kappa chain v r
27	26	17.6	27	2 S52355	hypothetical prote
28	26	17.6	28	2 S55442	beta A2 crystallin
29	25.5	17.2	17	2 I51203	myosin heavy chain

neuromedin C - lau
neuromedin C - bov
hypothetical prote
Ig mu chain V regi
gene c-fms protein
replication initia
homeodomain protei
thymidine phosphor
gastrin-releasing
aryldialkylphospha
monocyte chemotact
gastrin-releasing
gastrin-releasing
jacalin alpha chai
jacalin alpha chai
cell adhesion regu

30 25 16.9 10 2 PQ0177
31 25 16.9 10 2 A60647
32 25 16.9 12 2 C64030
33 25 16.9 14 2 A43847
34 25 16.9 16 2 I57330
35 25 16.9 23 2 B42382
36 25 16.9 23 2 S60569
37 25 16.9 24 2 S11384
38 25 16.9 25 2 S06263
39 25 16.9 25 2 PT0088
40 25 16.9 25 2 I67552
41 25 16.9 27 1 RHPGA
42 25 16.9 27 1 RHPGA
43 25 16.9 27 2 S29641
44 25 16.9 27 2 S29639
45 25 16.9 28 2 I54183

ALIGNMENTS

RESULT 1

RHPGG

gonadoliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Note: the synthetic and natural hormones have the same physicochemical and biologic

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.1%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 HWSYGLRP 27

Db 2 HWSYGLRP 9

RESULT 2

RHSHG

gonadoliberin - sheep

C:Species: Ovis orientalis arles, Ovis ammon aries (domestic sheep)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

A:Residues: 1-10 <BUR>
A>Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormones
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.1%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
DB 2 HWSYGLRP 9

RESULT 3
RHAQ1
gonadoliberein I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regal, P.; Pt. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 32.4%; Score 48; DB 1; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.67;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
DB 2 HWSYGLQP 9

RESULT 4
A21114
gonadoliberein - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match 26.4%; Score 39; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
DB 2 HWSYGLWP 9

RESULT 5
T01664
envelope protein - hepatitis C virus (fragment)
C:Species: hepatitis C virus

C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 17-Nov-2000
C:Accession: T01664
R:Scarselli, E.; Cerino, A.; Esposito, G.; Silini, E.; Mondelli, M.U.; Traboni, C.
J. Virol. 69, 4407-4412, 1995
A:Title: Occurrence of antibodies reactive with more than one variant of the putative A:Reference number: Z14388; MUID:95287497
A:Accession: T01664
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-27 <SCA>
A:Cross-references: EMBL:X79669; NID:g2276229; PIDN:CAA56117.1; PID:g2276230
C:Genetics:
A:Gene: E2/NS1
C:Superfamily: hepatitis C virus genome polypeptide

Query Match 23.6%; Score 35; DB 2; Length 27;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 FIGITELGPS 18
DB 16 FVGMESLGPS 25

RESULT 6
RHAQ2
gonadoliberein II - American alligator
N:Alternate names: gonadotropin-releasing hormone II
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: B60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regal, P.; Pt. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: B60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.0%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 73;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
DB 2 HWSHGWP 9

RESULT 7
A61126
gonadoliberein - spotted ratfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Hydrolagus colliei (spotted ratfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C:Accession: A61126
R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.
Gen. Comp. Endocrinol. 82, 152-161, 1991
A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a holobranch
A:Reference number: A61126; MUID:91340067
A:Accession: A61126
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.0%; Score 34; DB 1; Length 10;

Best Local Similarity 62.5%; Pred. NO. 73;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
|||:| |
Db 2 HWSHGWP 9

RESULT 8
A46030
gonadoliberin I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 23.0%; Score 34; DB 2; Length 10;
Best Local Similarity 62.5%; Pred. NO. 73;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
|||:| |
Db 2 HWSHGWP 9

RESULT 9
B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 23.0%; Score 34; DB 2; Length 10;
Best Local Similarity 62.5%; Pred. NO. 73;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
|||:| |
Db 2 HWSHGWP 9

RESULT 10
A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A:Title: Primary structure and biological activity of a third gonadotropin-releasing hor
A:Reference number: A49187; MUID:93178316
A:Accession: A49187

A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 20.9%; Score 31; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. NO. 2e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
|||:| |
Db 2 HWSHDWKP 9

RESULT 11
P00070
T-cell receptor beta chain (BTB15) - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
C:Accession: P00070
R:Tanaka, A.; Ishiguro, N.; Shinagawa, M.
submitted to JIPID, May 1990
A:Description: Sequence analysis of bovine T-cell receptor beta chain genes.
A:Reference number: JQ0472
A:Accession: P00070
A:Molecule type: mRNA
A:Residues: 1-22 <TAN>
A:Experimental source: T cell
C:Genetics:
A:Gene: BTB15
C:Keywords: receptor

Query Match 20.9%; Score 31; DB 2; Length 22;
Best Local Similarity 63.6%; Pred. NO. 4.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 14 ELGSLHWSYG 24
|||:| |
Db 1 ELGGA--WSYG 9

RESULT 12
PC4030
rRNA endonuclease (EC 3.1.27.10) - shiitake mushroom (fragment)
N:Alternate names: nuclease Le3
C:Species: Lentinula edodes (shiitake mushroom)
C:Date: 21-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 20-Mar-1996
C:Accession: PC4030
R:Kobayashi, H.; Inokuchi, N.; Koyama, T.; Tomita, M.; Irie, M.
Biosci. Biotechnol. Biochem. 59, 1169-1171, 1995
A:Title: Purification and characterization of the 2nd 5'-nucleotide-forming nuclease
A:Reference number: PC4030; MUID:95337563
A:Accession: PC4030
A:Molecule type: DNA
A:Residues: 1-23 <KOB>
C:Comment: This enzyme has 3'-nucleotidase activity.
C:Keywords: endonuclease; hydrolase

Query Match 19.6%; Score 29; DB 2; Length 23;
Best Local Similarity 46.2%; Pred. NO. 9.6e+02;
Matches 6; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 FIGITELGSLHW 21
|||:| |
Db 11 FIASNALDPSFVW 23

RESULT 13
S28436
major outer membrane protein V - Aeromonas hydrophilla (fragment)
C:Species: Aeromonas hydrophilla
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995

C:Accession: S28436
R:Jeanteur, D.; Gletsu, N.; Pattus, F.; Buckley, J.T.
Mol. Microbiol. 6, 3355-3363, 1992
A:Title: Purification of Aeromonas hydrophila major outer-membrane proteins: N-terminal
A:Reference number: S28433; MUID:93133115
A:Accession: S28436
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-21 <JEA>

Query Match 19.3%; Score 28.5; DB 2; Length 21;
Best Local Similarity 35.0%; Pred. NO. 1e+03;
Matches 7; Conservative 2; Mismatches 6; Indels 5; Gaps 1;

QY 5 ANSKFIGITELGSLHWSYG 24
|| : ||
Db 1 ANDWYGTG-----GAGWAYG 15

RESULT 14

S57517
T cell receptor beta chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57517
R:Burrows, S.R.; Sillins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Arguet, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57517
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-16 <BUR>
A:Cross-references: EMBL:249924; NID:9887480; PIDN:CAA90170.1; PID:9887481
C:Keywords: T-cell receptor

Query Match 18.9%; Score 28; DB 2; Length 16;
Best Local Similarity 55.6%; Pred. NO. 9.1e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 16 GPSLHWSYG 24
|| | : ||
Db 4 GPLRGNYG 12

RESULT 15

S57286
translation elongation factor aEF-1 beta - Sulfolobus solfataricus (fragments)
C:Species: Sulfolobus solfataricus
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S57286
R:Arcari, P.; Raimo, G.; Iannicciello, G.; Gallo, M.; Bocchini, V.
Biochim. Biophys. Acta 1263, 86-88, 1995
A:Title: The first nucleotide sequence of an archaeal elongation factor 1-beta gene.
A:Reference number: S57288; MUID:95359209
A:Accession: S57286
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <ARC>

Query Match 18.9%; Score 28; DB 2; Length 20;
Best Local Similarity 71.4%; Pred. NO. 1.2e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 IGITELG 16
: || | |
Db 13 VGITRLG 19

Search completed: October 10, 2002, 17:11:08
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:03:32 ; Search time 13 Seconds
(without alignments)
83.396 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKFIGITELGPSLHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 1673

Minimum DB seq length: 0

Maximum DB seq length: 28

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	48	32.4	10	GON1_ALLMI	P37041 alligator m
2	41	27.7	10	GON1_CLUPA	P81749 clupea pall
3	39	26.4	10	GON3_ONCKE	P20367 oncorhynch
4	34	23.0	10	GON2_CHICK	P37043 gallus gall
5	34	23.0	10	GON1_SQUAC	P27429 squalus aca
6	33	22.3	22	GSP2_PHYTA	P80661 physcomitre
7	31	20.9	10	GON3_PETNA	P30948 petromyzon
8	28	18.9	14	FIBA_HORSE	P14452 equus cabal
9	28	18.9	21	MDH_RHOPA	P80458 rhodopsedo
10	28	18.9	21	NRLA_ACTIS	P33036 acinetobact
11	28	18.9	28	GUN_SCHCO	P81190 schizophyll
12	27	18.2	18	SODM_MYCHA	P80582 mycobacteri
13	26.5	17.9	27	GRP_CHICK	P01295 gallus gall
14	26	17.6	10	GON1_CHEPR	P80677 chelyosoma
15	26	17.6	26	PCW4_PACGO	P82426 pachycondyl
16	25	16.9	8	ALL1_CYDPO	P82152 cydia pomon
17	25	16.9	10	GRP_RANRI	P23260 rana ridibu
18	25	16.9	22	CR33_LITCE	P56240 litoria cae
19	25	16.9	22	CR34_LITCE	P56241 litoria cae
20	25	16.9	23	GRP_ONCMY	Q9PS30 oncorhynch
21	25	16.9	23	TYPH_LACCA	P19663 lactobacill
22	25	16.9	25	GRP_SCYCA	P09472 scyllorhinu
23	25	16.9	27	GRP_CANFA	P08989 canis fami
24	25	16.9	27	GRP_PIG	P01294 sus scrofa
25	25	16.9	28	GRP_ALLMI	P31886 alligator m
26	24.5	16.6	8	AL16_CARMA	P81819 carcinus ma
27	24.5	16.6	20	SODM_HORVU	P28524 hordeum vul
28	24.5	16.6	27	CHI2_STROI	P29116 streptomyce
29	24	16.2	11	TKC2_CALVO	P41518 calliphora
30	24	16.2	13	SODM_CANFA	P54712 canis fami
31	24	16.2	13	TEML_RANTE	P57104 rana tempor
32	24	16.2	14	PSAG_CUCSA	P42049 cumicis sat
33	24	16.2	20	OMPI_ACTAC	P20242 actinobacil

ALIGNMENTS

RESULT 1

GON1_ALLMI
ID GON1_ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 32.4%; Score 48; DB 1; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.28; 0; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0;

Oy 20 HWSYGLRP 27

Db 2 HWSYGLQP 9
|||||:

RESULT 2

GON1_CLUPA
ID GON1_CLUPA STANDARD; PRT; 10 AA.
AC P81749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH)
DE (Luliberin I).
GN GNRH1.
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;

OC Clupeinae; Clupea.
OX NCBI_TaxID=30724;
RN [1]
RP SEQUENCE, AND FUNCTION
RC TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;

Query Match 27.7%; Score 41; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 3.1;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
| | | | |
Db 2 HWSYGLSP 9
| | | | |

RESULT 3
ID GON3_ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III) (LH-
RH III) (Luliberin III).
GN GNRH3.
OS Oncorhynchus keta (Chum salmon), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Proacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O.keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR; A21114; A21114.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

Query Match 26.4%; Score 39; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 6.1;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
| | | | |
Db 2 HWSYGLWP 9
| | | | |

RESULT 4
ID GON2_CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GNRH-II)
(LH-RH II) (Luliberin II).
OS Gallus gallus (Chicken).
OS Alligator mississippiensis (American alligator),
OS Squalus acanthias (Spiny dogfish),
OS Hydrolagus colliet (Spotted ratfish) (Pacific ratfish), and
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=Chicken; TISSUE=Hypothalamus;
RX MEDLINE=84222059; PubMed=642779;
RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
RA Matsuo H.;
RT "Identification of the second gonadotropin-releasing hormone in
chicken hypothalamus: evidence that gonadotropin secretion is
probably controlled by two distinct gonadotropin-releasing hormones
in avian species.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=A.mississippiensis; TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lauce V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
RN [3]
RP SEQUENCE.
RC SPECIES=S.acanthias; TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
dogfish brain provides insight into GNRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN [4]
RP SEQUENCE.
RC SPECIES=H.colliet; TISSUE=Brain;
RX MEDLINE=91340067; PubMed=1678723;
RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
Lee T.;
RT "Primary structure of gonadotropin-releasing hormone from the brain
of a holocephalan (ratfish: Hydrolagus colliet).";
RL Gen. Comp. Endocrinol. 82:152-161(1991).
RN [5]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

```
RA Chaig J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RT hormones, including a novel form, from an ancient teleost, herring.";
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: B60066; RHAQ2.
DR PIR: A61126; A61126.
DR PIR: B46030; B46030.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;

Query Match 23.0%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
| | | | |
DB 2 HWSHWYP 9

RESULT 5
GONL_SQUAC STANDARD; PRT; 10 AA.
AC P27429;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
DE (Luliberin).
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=923335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Namvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
RT dogfish brain provides insight into GNRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;

Query Match 23.0%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
| | | | |
DB 2 HWSHWLP 9

RESULT 6
PSP2_PHYPA STANDARD; PRT; 22 AA.
AC P80661;
```

```
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxygen-evolving enhancer protein 2 (OEE2) (22 kDa subunit of oxygen
DE evolving system of photosystem II) (fragments).
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBI_TaxID=3218;
RN [1]
RP SEQUENCE.
RC TISSUE=Protonema;
RX MEDLINE=97275459; PubMed=9129336;
RA Kasten B., Buck F., Nuske J., Reski R.;
RT "Cytokinin affects nuclear- and plastome-encoded energy-converting
RT plastid enzymes.";
RL Planta 201:261-272(1997).
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF PHOTOSYSTEM II.
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -!- INDUCTION: BY LIGHT.
CC -!- SIMILARITY: BELONGS TO THE PSBP FAMILY.
KW Photosynthesis; Photosystem II; Chloroplast; Thylakoid; Membrane;
KW Multigene family.
FT NON_CONS 15 16
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2316 MW; E1553CC575C54BB1 CRC64;

Query Match 22.3%; Score 33; DB 1; Length 22;
Best Local Similarity 53.8%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 ANSKFIGHTELGP 17
| : | | | | |
DB 10 AGNGFTXITEYGP 22

RESULT 7
GON3_PETMA STANDARD; PRT; 10 AA.
ID GON3_PETMA STANDARD; PRT; 10 AA.
AC P30948;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
DE (Luliberin III).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
RT releasing hormone from lamprey brain.";
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237AA1F5A3 CRC64;

Query Match 20.9%; Score 31; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 93;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 20 HWSYGLRP 27
    |||: :|
Db 2 HWSHDWKP 9

RESULT 8
FIBA_HORSE STANDARD; PRT; 14 AA.
AC P14452;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 14 FIBRINOPEPTIDE A.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1517 MW; 45998EB63C2A15E7 CRC64;

Query Match 18.9%; Score 28; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 13 TELGPSLHWSYGLR 26
    |||: :|
Db 1 TEEGFLHEGGVR 14

RESULT 9
MDH_RHOPA STANDARD; PRT; 21 AA.
AC P80458;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Malate dehydrogenase (EC 1.1.1.37) (Fragment).
GN MDH.
OS Rhodopsudomonas palustris.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Rhodopsudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE.
RA Naterstad K., Synstad B., Sirevag R.;
RL Submitted (SEP-1996) to the SWISS-PROT data bank.
CC -!- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR InterPro: IPR001252; MDH_actsite.
DR PROSITE: PS00068; MDH; PARTIAL.
KW Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2053 MW; FA97098FF856205F CRC64;

Query Match 18.9%; Score 28; DB 1; Length 21;
Best Local Similarity 45.5%; Pred. No. 5.4e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 IKANSKFITIGIT 13
    ||||: |
Db 1 VSYNSKFLAAT 11

RESULT 11
GUN_SCHCO STANDARD; PRT; 28 AA.
ID GUN_SCHCO
AC P81190;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Endoglucanase (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)
DE (Fragment).
OS Schizophyllum commune (Bracket fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Stereales; Schizophyllaceae; Schizophyllum.
OX NCBI_TaxID=5334;
RN [1]
RP SEQUENCE.
RA Clarke A.J., Drummelsmith J., Yaguchi M.;
RL "Identification of the catalytic nucleophile in the cellulase from
RT Schizophyllum commune and assignment of the enzyme to Family 5,
RT subtype 5 of the glycosidases.";
RL FEBS Lett. 414:359-361(1997).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
```



```

CC linkages in cellulose.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -!- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
DR InterPro: IPR001547; Glyco_hydro_F5.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; PARTIAL.
KW Cellulose degradation; Hydrolase; Glycosidase; Zymogen; Membrane;
KW Lipoprotein.
FT AC_SITE 20 20 NUCLEOPHILE.
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 2937 MW; B3F1C0C99C9950BE CRC64;
Query Match 18.9%; Score 28; DB 1; Length 28;
Best Local Similarity 38.9%; Pred. No. 7.1e+02;
Matches 7; Conservative 7; Mismatches 0; Indels 4; Gaps 2;
QY 1 QVIKANSK--FTGITELG 16
DB ::::: |:: |::
7 EWLKANNQRGFLG--EMG 22
RESULT 12
SODM_MYCHA STANDARD; PRT; 18 AA.
AC P80582;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn] (EC 1.15.1.1) (Fragment).
GN SODA OR SOD.
OS Mycobacterium habana.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1784;
RN [1]
RP SEQUENCE.
RC STRAIN=TMC 5135;
RX MEDLINE=96262709; PubMed=8704977;
RA Bisht D., Mehrotra J., Dhindsa M.S., Singh N.B., Sinha S.;
RT "A major T-cell-inducing cytosolic 23 kDa protein antigen of the
RT vaccine candidate Mycobacterium habana is superoxide dismutase.";
RL Microbiology 142:1375-1383(1996).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Manganese (By similarity).
CC -!- SUBUNIT: HOMODIMER (PROBABLE).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR HSP; P17670; 1IDS.
DR InterPro: IPR001189; SOD_MI.
DR Pfam: PF00081; sodfe; 1.
DR PROSITE: PS00088; SOD_MN; PARTIAL.
KW Oxidoreductase; Manganese.
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1883 MW; A4161A3DAC93F710 CRC64;
Query Match 18.2%; Score 27; DB 1; Length 18;
Best Local Similarity 57.1%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 17 PSLHWSY 23
DB | | | | |
6 PDLGWDY 12
RESULT 13
GRP_CHKCK STANDARD; PRT; 27 AA.
ID GRP_CHKCK
AC P01295;
DT 21-JUL-1986 (Rel. 01, Created)

```

```

DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Gastrin-releasing peptide (GRP) (Proventricular peptide) [Contains:
DE Neuromedin C (GRP-10)].
GN GRP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=81164953; PubMed=7215543;
RA McDonald T.J., Joernvall H., Ghatel M., Bloom S.R., Mutt V.;
RT "Characterization of an avian gastric (proventricular) peptide having
RT sequence homology with the porcine gastrin-releasing peptide and the
RT amphibian peptides bombesin and alaytesin.";
RL FEBS Lett. 122:45-48(1980).
CC -!- FUNCTION: GRP stimulates gastrin release as well as other
CC gastrointestinal hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
CC FAMILY.
DR PIR: A01563; RHCHA.
DR PIR: S08150; S08150.
DR InterPro: IPR000874; Bombesin.
DR Pfam: PF02044; Bombesin; 1.
DR PROSITE: PS00257; BOMBESIN; 1.
KW Bombesin family; Amidation.
FT PEPTIDE 18 27
FT MOD_RES 27 27
FT SEQUENCE 27 AA; 2842 MW; 3C121D0118D551C7 CRC64;
Query Match 17.9%; Score 26.5; DB 1; Length 27;
Best Local Similarity 35.7%; Pred. No. 1.1e+03;
Matches 5; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
QY 12 ITELGP-SLHWSYG 24
DB |::: |:: |
11 LTKIIPRGSHWVG 24
RESULT 14
GONL_CHEPR
ID GONL_CHEPR STANDARD; PRT; 10 AA.
AC P80677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I)
DE (Luliberin I)..
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
RT and the evolutionary implications";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.

```

CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
 CC THE GONADS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
 CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
 CC -!- MASS SPECTROMETRY: MW=1246.56; METHOD=MALDI.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.

DR InterPro: IPR002012; GNRH.
 DR PROSITE: PS004773; GNRH; 1.
 DR Hormone; Amidation.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
 SQ SEQUENCE 10 AA; 1264 MW; 284B3639DB5AB5A3 CRC64;

Query Match 17.6%; Score 26; DB 1; Length 10;
 Best Local Similarity 50.0%; Pred. NO. 5.1e+02;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
 Db 2 HWSDFPKP 9

RESULT 15

PCW4_PACGO
 ID PCW4_PACGO STANDARD; PRT; 26 AA.
 AC P82426;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Ponericin W4.
 OS Pachycondyla goeldii (Ponerine ant.).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
 OC Formicidae; Ponerinae; Pachycondyla.
 OX NCBI_TaxID=118888;
 RN [1]
 RP SEQUENCE, AND FUNCTION.
 RC TISSUE-Venom;
 RX PubMed=11279030;
 RA Orivel J., Redeker V., Le Caer J.-P., Krier F., Revol-Junelles A.-M.,
 RA Longeon A., Chafotte A., Dejean A., Rossier J.;
 RT "Ponericins, new antibacterial and insecticidal peptides from the
 RT venom of the ant Pachycondyla goeldii.";
 RL J. Biol. Chem. 276:17823-17829(2001).
 CC -!- FUNCTION: BROAD SPECTRUM OF ACTIVITY AGAINST BOTH GRAM-POSITIVE
 CC AND GRAM-NEGATIVE BACTERIA AND S.CEREVISIAE. HAS INSECTICIDAL
 CC AND HEMOLYTIC ACTIVITIES.
 CC -!- MASS SPECTROMETRY: MW=2851.81; METHOD=MALDI.
 KW Antibiotic; Insect immunity; Fungicide; Hemolysis.
 SQ SEQUENCE 26 AA; 2853 MW; E361FBB1F5D4F70F CRC64;

Query Match 17.6%; Score 26; DB 1; Length 26;
 Best Local Similarity 41.7%; Pred. NO. 1.3e+03;
 Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 16 GPSLHWSYGLRP 27
 Db 4 GTALKWGVKLLP 15

Search completed: October 10, 2002, 17:09:41
 Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:04:12 ; Search time 29 seconds
(without alignments)
167.030 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKFIGITELGSLHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 12285

Minimum DB seq length: 0

Maximum DB seq length: 28

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	40	27.0	24	2 Q56150	Q56150 streptomyces
2	36	24.3	27	12 Q68906	Q68906 hepatitis c
3	35	23.6	27	12 Q39931	Q39931 hepatitis c
4	35	23.6	27	12 Q39932	Q39932 hepatitis c
5	34	23.0	18	2 Q9R4C3	Q9R4C3 agrobacteri
6	33	22.3	27	12 Q9I1K5	Q9I1K5 hepatitis c
7	32.5	22.0	27	12 Q9I1P8	Q9I1P8 hepatitis c
8	32	21.6	20	2 P97L46	P97L46 escherichia
9	32	21.6	25	4 Q9UD88	Q9UD88 homo sapien
10	32	21.6	27	12 Q9Q9F2	Q9Q9F2 hepatitis c
11	32	21.6	27	12 Q9Q9E6	Q9Q9E6 hepatitis c
12	32	21.6	27	12 Q9Q9E2	Q9Q9E2 hepatitis c
13	32	21.6	27	12 Q9Q9D6	Q9Q9D6 hepatitis c
14	31	20.9	27	12 Q37180	Q37180 hepatitis c
15	31	20.9	27	12 Q9Q1E7	Q9Q1E7 hepatitis c
16	31	20.9	27	12 Q9Q1E5	Q9Q1E5 hepatitis c

ALIGNMENTS

RESULT 1

Q56150 ID Q56150 PRELIMINARY; PRT; 24 AA.
AC Q56150;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 2.4 KDA PROTEIN (FRAGMENT).
OS Streptomyces violaceus (Streptomyces venezuelae).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1936;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISP5230;
RX MEDLINE=96070798; PubMed=7592948;
RA Mosher R.H., Camp D.J., Yang K., Brown M.P., Shaw W.V., Vining L.C.;
RT "Inactivation of chloramphenicol by O-phosphorylation: A novel
RT resistance mechanism in Streptomyces venezuelae ISP5230, a
RT chloramphenicol producer.";
RL J. Biol. Chem. 270:27000-27006(1995).
DR EMBL; U09991; AAB36571.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2443 MW; 49DDCD3DB5FF0745 CRC64;

Query Match 27.0%; Score 40; DB 2; Length 24;

Best Local Similarity 63.6%; Pred. No. 29;

Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 ELGPSLHWSYG 24

|||||

Db 1 ELGPDHPHWVCG 11

RESULT 2

Q68906

ID Q68906 PRELIMINARY; PRT; 27 AA.

AC Q68906;

```
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE (NS1)] (FRAGMENT).
GN E2/NS1.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GENOTYPE II;
RX MEDLINE=95287497; PubMed=7539508;
RA Scarselli E., Cerino A., Esposito G., Silini E., Mondelli M.U.,
RA Traboni C.;
RT "Occurrence of antibodies reactive with more than one variant of the
RT putative envelope glycoprotein (gp70) hypervariable region 1 in
RT viremic hepatitis C virus infected patients.";
RL J. Virol. 69:4407-4412(1995).
DR EMBL; X79671; CAA56119.1; -
DR InterPro: IPR002531; HCV.NS1.
DR Pfam: PF01560; HCV.NS1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2782 MW; BCFEDC0FC350AD4 CRC64;

Query Match 24.3%; Score 36; DB 12; Length 27;
Best Local Similarity 46.7%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 KANSKFIFGITELGSPS 18
:|:|:|:|
DB 11 RATSGFVGMFSPGPS 25

RESULT 3
O39931 PRELIMINARY; PRT; 27 AA.
AC O39931;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE (NS1)] (FRAGMENT).
GN E2/NS1.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GENOTYPE II;
RX MEDLINE=95287497; PubMed=7539508;
RA Scarselli E., Cerino A., Esposito G., Silini E., Mondelli M.U.,
RA Traboni C.;
RT "Occurrence of antibodies reactive with more than one variant of the
RT putative envelope glycoprotein (gp70) hypervariable region 1 in
RT viremic hepatitis C virus infected patients.";
RL J. Virol. 69:4407-4412(1995).
DR EMBL; X79669; CAA56117.1; -
DR InterPro: IPR002531; HCV.NS1.
DR Pfam: PF01560; HCV.NS1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2878 MW; AD948DC0FC240FC4 CRC64;

Query Match 23.6%; Score 35; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 FFIGITELGSPS 18
:|:|:|:|
DB 16 FVGMFSLGSPS 25

RESULT 4
O39932 PRELIMINARY; PRT; 27 AA.
AC O39932;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
DE (NS1)] (FRAGMENT).
GN E2/NS1.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GENOTYPE II;
RX MEDLINE=95287497; PubMed=7539508;
RA Scarselli E., Cerino A., Esposito G., Silini E., Mondelli M.U.,
RA Traboni C.;
RT "Occurrence of antibodies reactive with more than one variant of the
RT putative envelope glycoprotein (gp70) hypervariable region 1 in
RT viremic hepatitis C virus infected patients.";
RL J. Virol. 69:4407-4412(1995).
DR EMBL; X79670; CAA56118.1; -
DR InterPro: IPR002531; HCV.NS1.
DR Pfam: PF01560; HCV.NS1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2839 MW; BCFEDC0ED6B23D4 CRC64;

Query Match 23.6%; Score 35; DB 12; Length 27;
Best Local Similarity 40.0%; Pred. No. 2e+02;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 KANSKFIFGITELGSPS 18
:|:|:|:|
DB 11 RATNSFVGMFSPGPS 25

RESULT 5
O3994C3 PRELIMINARY; PRT; 18 AA.
AC O3994C3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROTOCATCHUATE 3,4-DIOXYGENASE TYPE I BETA SUBUNIT (EC 1.13.11.3)
DE (FRAGMENT).
OS Agrobacterium tumefaciens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE.
RX MEDLINE=96337865; PubMed=8772173;
RA Hammer A., Stolz A., Knackmuss H.;
RT "Purification and characterization of a novel type of protocatechuate
RT 3,4-dioxygenase with the ability to oxidize 4-sulfocatechol.";
RL Arch. Microbiol. 166:92-100(1996).
SQ SEQUENCE 18 AA; 2008 MW; C7EBD971BBE5BBE9 CRC64;

Query Match 23.0%; Score 34; DB 2; Length 18;
Best Local Similarity 41.7%; Pred. No. 1.8e+02;
```

Matches, 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 14 ELGPSLHWSYGL 25
| | | | |
Db 6 ETGPTFTWDXGI 17

RESULT 6
Q9IIK5 PRELIMINARY; PRT; 27 AA.
AC Q9IIK5; 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70) (NS1)] (FRAGMENT).
DE Hepatitis C virus.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RA Alberto S.-F.;
RT "Influence of the dynamics of Hepatitis C virus quasiespecies in the histological outcome of liver transplantation."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF221440; AAF78002.1; -
DR InterPro: IPR002531; HCV_NSI.
DR Pfam: PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2879 MW; F703C29224423E73 CRC64;

Query Match 22.3%; Score 33; DB 12; Length 27;
Best Local Similarity 41.7%; Pred. No. 4e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 SKFIGITELGPS 18
: | : | | | |
Db 13 NRVSLFSLGPS 24

RESULT 7
Q9IIP8 PRELIMINARY; PRT; 27 AA.
AC Q9IIP8; 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RA Alberto S.-F.;
RT "Influence of the dynamics of Hepatitis C virus quasiespecies in the histological outcome of liver transplantation."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF221397; AAF77959.1; -
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2987 MW; 04A47C9B4D26C7C2 CRC64;

Query Match 22.0%; Score 32.5; DB 12; Length 27;
Best Local Similarity 53.8%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 15 LGFSLHWSYGLRP 27
| | | | |

Db 15 LHPSEHWG---RP 24

RESULT 8
P97146 PRELIMINARY; PRT; 20 AA.
ID P97146; 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE BETA-LACTAMASE (FRAGMENT).
GN BLA.
OS Escherichia coli.
OG Plasmid.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Morales V.M., Backman A., Bagdasarjan M.;
RT "Sequence of junctions of expression vector pMMB666EH."
RL Nucleic Acids Res. 0:0-0(0).
DR EMBL: X15193; CAA33267.1; -
DR HSSP: P00810; 1BTL.
KW Plasmid.
FT NON_TER 1
SQ SEQUENCE 20 AA; 2369 MW; 8FF2DD7208E77B69 CRC64;

Query Match 21.6%; Score 32; DB 2; Length 20;
Best Local Similarity 58.3%; Pred. No. 4.1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 12 ITELGPSSL--HW 21
| | | | |
Db 9 IAEIGASLIKHW 20

RESULT 9
Q9UD88 PRELIMINARY; PRT; 25 AA.
ID Q9UD88; 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE APO(A) KRINGLE 4-37, APO(A) KRINGLE TYPE 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94060120; PubMed=7848387;
RA Pfaffinger D., McLean J., Scanu A.M.;
RT "Amplification of human APO(a) kringle 4-37 from blood lymphocyte DNA."
RL Biochim. Biophys. Acta 1225:107-109(1993).
DR HSSP: P00747; 1KRN.
SQ SEQUENCE 25 AA; 2928 MW; BEEDD4C62F480A8 CRC64;

Query Match 21.6%; Score 32; DB 4; Length 25;
Best Local Similarity 44.4%; Pred. No. 5.3e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 15 LGFSLHWSY 23
: | | | |
Db 4 MDPISIRWEY 12

RESULT 10
Q9Q9F2 PRELIMINARY; PRT; 27 AA.
ID Q9Q9F2
AC Q9Q9F2;

Df 11 RATHSFVSIFSTGSPS 25

RESULT 12

Q9Q9E2 PRELIMINARY; PRT; 27 AA.

ID Q9Q9E2 AC Q9Q9E2 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70) (NS1)] (FRAGMENT).

DE DE Hepatitis C virus.

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

NCBI_TaxID=11103;

RN [1]

RX MEDLINE=20091320; PubMed=10623727;

RA Sandres K., Dubois M., Pasquier C., Payen J.L., Atric L., Duffaut M.

RA Vinel J.P., Pascal J.P., Pucl J., Izopet J.;

RT "Genetic heterogeneity of hypervariable region 1 of the hepatitis C virus (HCV) genome and sensitivity of HCV to alpha interferon therapy.";

RL J. Virol. 74: 661-668(2000).

RE EMBL; AF166558; AAD52216.1; -

DR InterPro: IPR002531; HCV_NSL1.

DR Pfam: PF01560; HCV_NSL1; 1.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;

KW Polyprotein; Transmembrane.

FT NON_TER 1

FT NON_TER 27

SQ SEQUENCE 27 AA; 2735 MW; 35260421E0DA78D7 CRC64;

Query Match 21.6%; Score 32; DB 12; Length 27;
Best Local Similarity 40.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps

QY 4 KANSKFGITELGPS 18
:-|:-|||

Df 11 RATHSFVSIFSTGSPS 25

RESULT 13

Q9Q9D6 PRELIMINARY; PRT; 27 AA.

ID Q9Q9D6 AC Q9Q9D6 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70) (NS1)] (FRAGMENT).

DE DE Hepatitis C virus.

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

NCBI_TaxID=11103;

RN [1]

RX MEDLINE=20091320; PubMed=10623727;

RA Sandres K., Dubois M., Pasquier C., Payen J.L., Atric L., Duffaut M.

RA Vinel J.P., Pascal J.P., Pucl J., Izopet J.;

RT "Genetic heterogeneity of hypervariable region 1 of the hepatitis C virus (HCV) genome and sensitivity of HCV to alpha interferon therapy.";

RL J. Virol. 74: 661-668(2000).

RE EMBL; AF166564; AAD52222.1; -

DR InterPro: IPR002531; HCV_NSL1.

DR Pfam: PF01560; HCV_NSL1; 1.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;

KW Polyprotein; Transmembrane.

FT NON_TER 1

FT NON_TER 27

SQ SEQUENCE 27 AA; 2735 MW; 35260421E0DA78D7 CRC64;

Query Match 21.6%; Score 32; DB 12; Length 27;
Best Local Similarity 40.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 KANSKFIGITELGSPS 18
: | | | | |
DB 11 RATHSFVSIFSTGSPS 25

RESULT 14

O37180 PRELIMINARY; PRT; 27 AA.
AC O37180;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ENVELOPE PROTEIN 2 (FRAGMENT).
GN E2.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=10;
RX MEDLINE=98105815; PubMed=9445070;
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
RT "Genetic diversity and tissue compartmentalization of the hepatitis C
virus genome in blood mononuclear cells, liver, and serum from chronic
hepatitis C patients.";
RL J. Virol. 72:1640-1646(1998).
DR EMBL: AF018289; AAC03570.1; -.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2600 MW; 9B84D51E2E58BBD8 CRC64;

Query Match 20.9%; Score 31; DB 12; Length 27;
Best Local Similarity 46.2%; Pred. No. 8.2e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 KANSKFIGITELG 16
: | | | | |
DB 11 QATSKFVGSSTQG 23

RESULT 15

Q9QIE7 PRELIMINARY; PRT; 27 AA.
AC Q9QIE7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)
(NS1)] (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.;
RT "The genetic heterogeneity of hypervariable region 1 of the viral
genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF166630; AAD52288.1; -.
DR InterPro: IPR002531; HCV_NS1.
DR Pfam: PF01560; HCV_NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
Polyprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 27 27

SQ SEQUENCE 27 AA; 2727 MW; 8CA75E43B6EB5B61 CRC64;
Query Match 20.9%; Score 31; DB 12; Length 27;
Best Local Similarity 54.5%; Pred. No. 8.2e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 7 SKFIGITELGP 17
: | | | | |
DB 14 SAFAGLFRLLGP 24

Search completed: October 10, 2002, 17:10:30
Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: October 10, 2002, 17:02:47 ; Search time 27 Seconds
(without alignments)
115.188 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKFGITELGSLHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 268091

Minimum DB seq length: 0

Maximum DB seq length: 28

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	112.5	76.0	27	15	AA62701	LHRH-containing im
2	74	50.0	15	11	AA06310	Tetanus toxin epit
3	74	50.0	15	18	AAW3506	Universal T-cell e
4	74	50.0	15	18	AAW11505	Tetanus toxoid uni
5	74	50.0	15	19	AAW67033	Tetanus toxin frag
6	74	50.0	15	19	AAW71321	Universal helper T
7	74	50.0	15	20	AAW04051	T-helper epitope f
8	74	50.0	15	20	AAW67578	T-cell epitope pep
9	74	50.0	15	20	AAW73220	Tetanus toxoid epi
10	74	50.0	15	21	AAW45511	Tetanus p2 epitope
11	74	50.0	15	21	AAW82637	Tetanus toxoid T c

12	74	50.0	15	21	AA92625	Foreign epitope P2
13	74	50.0	15	21	AA984427	Amino acid sequenc
14	74	50.0	15	21	AAW70300	Clostridium tetani
15	74	50.0	15	21	AAW44763	Tetanus toxoid pro
16	74	50.0	15	22	AAEL1763	Clostridium tetani
17	74	50.0	15	22	AAW99515	Vaccine related MH
18	74	50.0	15	22	AAW85701	Amino acid sequenc
19	74	50.0	15	22	AAW85451	Wild-type TT830 (t
20	74	50.0	15	22	AAW61956	Tetanus Toxoid uni
21	74	50.0	15	22	AAW20143	Tetanus toxin T-ce
22	74	50.0	15	22	AAW68636	HER-2 B cell pepti
23	74	50.0	15	22	AAW46172	Tetanus toxoid TT8
24	74	50.0	15	22	AAW49071	Tetanus toxoid TT
25	74	50.0	16	18	AAW35445	T-cell stimulatory
26	74	50.0	16	20	AAW29705	Clostridium tetani
27	74	50.0	17	15	AAW62692	Helper T cell epit
28	74	50.0	17	15	AAW82573	Tetanus toxin help
29	74	50.0	17	17	AAW05599	Tetanus toxin help
30	74	50.0	17	17	AAW88395	T-cell antigen TT2
31	74	50.0	17	21	AAW99274	HLA class II bindi
32	74	50.0	17	21	AAW80056	Pathogen derived T
33	74	50.0	17	21	AAW54539	T helper cell (Th)
34	74	50.0	17	21	AAW58768	Unidentified pepti
35	74	50.0	17	22	AAW99516	Vaccine related MH
36	74	50.0	17	22	AAW62904	Amino acid residue
37	74	50.0	17	22	AAW84435	Amino acid sequenc
38	74	50.0	17	22	AAW30941	Amino acid sequenc
39	74	50.0	17	22	AAW31029	Antigenic fragment
40	74	50.0	17	22	AAW31118	Antigenic fragment
41	74	50.0	17	22	AAW15589	Peptide 5 for pept
42	74	50.0	18	20	AAW26607	HIV-derived lipope
43	74	50.0	19	21	AAW99055	HLA class II bindi
44	74	50.0	19	22	AAW99517	Vaccine related MH
45	74	50.0	22	22	AAW46175	Tetanus toxoid 830

ALIGNMENTS

```

RESULT 1
AA62701
ID  AA62701 standard; peptide; 27 AA.
XX
AC  AA62701;
XX
DT  10-SEP-1995 (first entry)
XX
DE  LHRH-containing immunogenic peptide.
XX
KW  Helper T cell epitope; universal immune stimulator; invasive; hapten;
KW  vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW  androgen-dependent carcinoma; antitumour; infertility;
KW  tetanus toxin.
XX
OS  Synthetic.
XX
FH  Key
FT  Domain
FT  Domain
FT  Domain
FT  Domain
XX
PN  WO9425060-A.
XX
PD  10-NOV-1994.
XX
PF  28-APR-1994; 9AWO-US04832.
XX
PR  27-APR-1993; 93US-0057166.
XX
PR  14-APR-1994; 94US-0229275.
XX
PA  (LADD/) LADD A E.
PA  (WANG/) WANG C Y.

```

```

Location/Qualifiers
1..17
/note= "tetanus toxin helper T cell epitope"
18..27
/note= "LHRH hapten"

```


CC binding activities can be used for the promotion of cell-attachment to
 CC plastic surfaces, in particular to inhibit tumour growth and metastasis,
 CC and for promotion of wound healing. Also a derivatised (A) can be used
 CC for the selection of specifically-binding aptamers or as a diagnostic
 CC agent. Such diagnostic-(A) molecules could be used to detect molecules
 CC derived from or indicative of pregnancy or of a disease, such as an
 CC infectious, autoimmune or cancerous disease.

XX Sequence 15 AA;

Query Match 50.0%; Score 74; DB 18; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIRANSKFIGITEL 15

Db 1 QYIRANSKFIGITEL 15

RESULT 4

AAW11505
 ID AAW11505 standard; Protein; 15 AA.

XX

AC AAW11505;

XX 24-SEP-1997 (first entry)

DT Tetanus toxoid universal Th epitope TT830.

DE Humanised antibody; anti-Fc receptor; H22; bifunctional; bispecific;
 XX fusion protein; chimera; tetanus toxoid; helper T cell epitope;
 KW antigen presentation; ds.
 KW Clostridium tetani.

XX WO9640789-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US09988.

XX 07-JUN-1995; 95US-0484172.

XX (MEDA-) MEDAREX INC.

XX Deco YM, Goldstein J, Graziano R, Somasundaram C;

XX WPI; 1997-052242/05.

XX N-PSDB; AAT58127.

XX Recombinant, multi-specific anti-Fc receptor antibody molecules -
 PT also comprise an anti-target portion, used for the treatment of
 PT cancer, autoimmune disease and pathogenic infection

XX Example 7; Fig 24; 115pp; English.

XX Synthetic DNA coding for the wild-type universal Th epitope from
 CC tetanus toxoid, designated TT830, was fused to the 3'-end of DNA
 CC encoding heavy chain sequences from the humanised anti-Fc gamma RI
 CC monoclonal antibody H22. The resulting fusion protein was shown to
 CC be significantly more efficient in antigen presentation and T cell
 CC stimulation than the TT830 epitope alone. A similar fusion
 CC construct was prepared coding for a mutant, antagonistic form of the
 CC epitope (designated TT833S) fused to the anti-Fc gamma RI. The
 CC Fab22-TT833S is at least 100 times more effective than TT833S in
 CC inhibiting T cell activation.

XX Sequence 15 AA;

Query Match 50.0%; Score 74; DB 18; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIRANSKFIGITEL 15

Db 1 QYIRANSKFIGITEL 15

RESULT 5

AAW67033
 ID AAW67033 standard; peptide; 15 AA.

XX

AC AAW67033;

XX 15-DEC-1998 (first entry)

DT Tetanus toxin fragment (residues 830-844).

DE Tetanus toxin; vaccine; antibody; carbohydrate peptide conjugate;
 KW dendrimeric poly-lysine; epitope; tumour.

XX Clostridium tetani.

XX WO9843677-A1.

XX 08-OCT-1998.

XX 27-MAR-1998; 98WO-EP01922.

XX 27-MAR-1997; 97US-0041726.

XX (INSP) INST PASTEUR.

XX Bay S, Cantacuzene D, Leclerc C, Lo-man R;

XX WPI; 1998-557071/47.

XX Carbohydrate peptide conjugate used as vaccine - comprises carrier
 PT with dendrimeric poly-lysine enabling multiple epitopes to be
 PT covalently attached

XX Disclosure; Page 13; 55pp; English.

XX The invention relates to a new carbohydrate peptide conjugate, which
 CC comprises a carrier with a dendrimeric poly-lysine enabling multiple
 CC epitopes to be covalently attached to it. Also claimed are: (1) an
 CC antibody purified from biological fluid or cells of organisms
 CC administered with the carbohydrate peptide conjugate, and (2) a diagnosis
 CC kit comprising antigen-specific antibodies elicited by immunisation with
 CC the carbohydrate peptide conjugate. The peptide conjugate, antibody and
 CC diagnosis kit are used to provide pharmaceutical compositions and
 CC vaccines against tumours. These can be used to support an immune response
 CC against viral infections caused by hepatitis virus, HIV or cytomegalo
 CC virus. They can be used to enhance immune responses, especially B- and T-
 CC cell responses, of humans and animals against bacterial infections. The
 CC carbohydrate peptide conjugate stimulates the antibody and T-cell
 CC response without stimulating undesired immune responses. The composition
 CC is capable of increasing the survival of tumour bearing humans and
 CC animals. The present sequence corresponds to residues 830-844 of tetanus
 CC toxin. The synthetic peptide corresponding to this sequence may be used
 CC as an epitope in a carbohydrate peptide conjugate.

XX Sequence 15 AA;

Query Match 50.0%; Score 74; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIRANSKFIGITEL 15

Db 1 QYIRANSKFIGITEL 15

RESULT 6

AAW71321
 ID AAW71321 standard; peptide; 15 AA.

XX AAW71321;
 XX 26-NOV-1998 (first entry)
 XX Universal helper T-cell epitope P2 derived from tetanus toxin.
 DE
 XX Liver stage; Plasmodium; Navy Yoelii Liver Stage 3 antibody; NYLS3;
 KW hepatic and erythrocytic stage protein; PyHEP17; vaccine;
 KW malaria parasite; teanus toxin; P2; helper T-cell epitope.
 XX Synthetic.
 OS Clostridium tetani.
 XX US5814617-A.
 XX 29-SEP-1998.
 XX 07-OCT-1994; 94US-0319704.
 XX 07-OCT-1994; 94US-0319704.
 XX (USNA) US SEC OF NAVY.
 XX Charoenvit Y, Doolan DL, Hedstrom RC, Hoffman SL;
 PI WPI; 1998-541794/46.
 DR
 XX Vaccine for protecting mammal against infection by malaria caused by
 PT Plasmodium species - comprises a first nucleic acid encoding a first
 PT polypeptide capable of eliciting an immune reaction against an
 PT antigen expressed during the liver
 XX
 PS Disclosure: Column 12; 24pp; English.
 XX AAW71321-22 represent universal helper T-cell epitopes derived from
 CC tetanus toxin. They are used to enhance host immune response to
 CC vaccines. The specification describes a Plasmodium yoelii liver stage
 CC 17 kDa hepatic and erythrocytic stage protein designated PyHEP17. This
 CC protein elicits a response from an Ig1 monoclonal antibody designated
 CC Navy Yoelii Liver Stage 3 (NYLS3). This antibody does not recognise
 CC sporozoites, but does recognise P. yoelii liver stage parasites. NYLS3
 CC eliminates upto 90% of liver stage parasites. The specification
 CC describes a vaccine for reducing the severity or incidence of infection
 CC by a malaria parasite of the genus Plasmodium. The DNA vaccine comprises
 CC exon 1 and part of exon 2 of the PyHEP17 gene.
 XX
 SQ Sequence 15 AA;
 Query Match 50.0%; Score 74; DB 19; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITEL 15
 DB 1 QYIKANSKFIGITEL 15
 RESULT 7
 ID AAY04051 standard; peptide; 15 AA.
 XX AAY04051;
 XX 04-JAN-2000 (first entry)
 XX T-helper epitope from tetanus toxoid.
 DE
 XX Covalently reactive antigen analog; CRAA; catalytic antibody;
 KW electrophilic reaction centre; phosphonate; boronate; vaccine;
 KW transition state analog; TSA; isostere; gp120; HIV-1; T-helper;
 KW tetanus; toxoid; B-T-epitope.
 XX

OS Clostridium tetani.
 XX WO9948925-A1.
 XX 30-SEP-1999.
 XX 23-MAR-1999; 99WO-US06325.
 XX 23-MAR-1998; 98US-0046373.
 XX (UYNE-) UNIV NEBRASKA.
 XX Paul S, Gololobov G, Smith L;
 PI WPI; 1999-591076/50.
 DR
 XX New covalently reactive antigen analogs used for treating e.g.
 PT autoimmune diseases, lymphoproliferative disorders, cancers, microbial
 PT infections, ischemic and reperfusion injury or septic shock -
 XX
 PS Disclosure: Page 86; 158pp; English.
 XX The patent discloses new covalently reactive antigen analogs (CRAA)
 CC of formula XI-Y-E-X2, in which XI and X2 represent peptide sequences
 CC of an epitope of a disease-associated protein, Y is a positively
 CC charged amino acid residue, preferably Lys or Arg, and E is an
 CC electrophilic reaction centre, preferably a phosphonate or boronate
 CC moiety. Depending on the identity of the epitope, the CRAA may be used
 CC to stimulate production of catalytic antibodies specific for
 CC predetermined antigens associated with particular medical disorders.
 CC They may also be used to permanently inactivate endogenously produced
 CC catalytic antibodies produced in certain autoimmune diseases as well as
 CC in certain lymphoproliferative disorders.
 CC Amongst the specifically exemplified CRAAs is one based on residues
 CC 421-436 of a B-cell epitope of gp120 (see AAY04046) which may be used
 CC to counter HIV-1 infections. When used as an immunogen, preferably this
 CC CRAA is conjugated at its N-terminal to a T-helper epitope from tetanus
 CC toxoid. The present sequence represents the T-helper epitope and
 CC corresponds to residues 830-844 of the toxoid.
 XX
 SQ Sequence 15 AA;
 Query Match 50.0%; Score 74; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITEL 15
 DB 1 QYIKANSKFIGITEL 15
 RESULT 8
 ID AAW67578 standard; peptide; 15 AA.
 XX AAW67578;
 XX 02-MAR-1999 (first entry)
 XX T-cell epitope peptide #4 for chimeric fimbria/T-cell epitope peptide.
 DE
 XX Chimeric; non-typable Haemophilus influenzae; fimbria; T-cell epitope;
 KW immunogenic composition; immune response.
 XX Synthetic.
 XX US5843464-A.
 XX 01-DEC-1998.
 XX 02-JUN-1995; 95US-0460502.
 XX 02-JUN-1995; 95US-0460502.

XX (OHIS) UNIV OHIO STATE.
 XX Bakaletz LO, Kaumaya PTP;
 XX WPI; 1999-044514/04.
 XX Synthetic chimeric fimbria peptide - useful for vaccination against
 XX non-typable Haemophilus influenzae
 XX Disclosure; Column 4; 16pp; English.
 XX The invention relates to the manufacture of a synthetic chimeric peptide
 XX comprising a non-typable Haemophilus influenzae fimbria peptide fused via
 XX a linker peptide to a T-cell epitope peptide. The chimeric peptide is
 XX used in immunogenic compositions which induce an immune response against
 XX non-typable Haemophilus influenzae. This sequence represents an example
 XX of a T-cell epitope peptide used to generate the chimeric peptide.
 XX Sequence 15 AA;
 SQ Query Match 50.0%; Score 74; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITEL 15
 DB 1 QYIKANSKFIGITEL 15
 RESULT 9
 AAW73220
 ID AAW73220 standard; Protein; 15 AA.
 XX AC AAW73220;
 XX DT 25-JAN-1999 (first entry)
 XX DE Tetanus toxoid epitope.
 XX KW Multispecific single chain antibody; antibody H22; tumour cell; therapy;
 XX antibody-dependent cellular cytotoxicity; ADCC; HER 2/neu; infection;
 XX epidermal growth factor receptor; breast cancer; ovarian cancer.
 XX OS Synthetic.
 XX PN US5837243-A.
 XX PD 17-NOV-1998.
 XX PF 07-JUN-1996; 96US-0661052.
 XX PR 07-JUN-1996; 96US-0661052.
 XX PR 07-JUN-1995; 95US-0484172.
 XX PA (MEDA-) MEDAREX INC.
 XX PI Deo YM, Goldstein J, Graziano R, Somasundaram C;
 XX WPI; 1999-023374/02.
 XX Specific killing of tumour cells - using a multi-specific molecule
 XX comprising an anti-Fc receptor antibody and a portion which binds to
 XX a target cell
 XX Example 7; Column 27; 57pp; English.
 XX This sequence represents a tetanus toxoid epitope and is recognised
 XX by the multispecific single chain antibody designated H22. The
 XX antibody can be used in the method of the invention for inducing
 XX antibody-dependent cellular cytotoxicity (ADCC) against a tumour cell
 XX which is characterised by overexpression of HER 2/neu or epidermal growth
 XX factor receptor (EGFR), comprises contacting the tumour cell with a

CC multispecific protein molecule (preferably a single chain antibody)
 CC comprising: (a) an anti-Fc receptor antibody or an antigen binding
 CC fragment; (b) a portion which binds to HER 2/neu; and (c) a portion which
 CC binds to EGFR. The method can be used for treating cancers especially
 CC breast cancer or ovarian cancer. The multispecific antibody can also
 CC be administered prophylactically to vaccinate a subject against infection
 CC by a target cell.
 XX Sequence 15 AA;
 SQ Query Match 50.0%; Score 74; DB 20; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITEL 15
 DB 1 QYIKANSKFIGITEL 15
 RESULT 10
 AAB45511
 ID AAB45511 standard; Protein; 15 AA.
 XX AC AAB45511;
 XX DT 26-FEB-2001 (first entry)
 XX DE Tetanus P2 epitope SEQ ID NO: 23.
 XX KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
 XX cancer; eosinophilia; vaccine; allergic rhinitis.
 XX OS Clostridium tetani.
 XX PN WO200065058-A1.
 XX PD 02-NOV-2000.
 XX PF 19-APR-2000; 2000WO-DK00205.
 XX PR 23-APR-1999; 99DK-0000552.
 XX PR 06-MAY-1999; 99US-0132811.
 XX PA (MEBI-) M & E BIOTECH AS.
 XX PI Klysner S;
 XX WPI; 2000-672791/65.
 XX Down-regulating interleukin 5 (IL-5) activity in humans by
 XX administering IL-5 and/or an IL-5 analogue, useful in the treatment,
 XX prophylaxis or amelioration of asthma or other chronic allergic
 XX conditions -
 XX Example 1; Page 137; 172pp; English.
 XX The present invention is concerned with methods of treating asthma,
 XX eosinophilia, allergic rhinitis and other allergic diseases. These
 XX involve the use of interleukin-5 (IL-5) analogues and modified IL-5
 XX proteins and their coding sequences to down-regulate IL-5 activity and
 XX thus reduce eosinophil numbers. The allergic diseases may be treated
 XX using autovaccines, nucleic acid vaccines or live vaccines. In addition,
 XX it is possible that they may be used in the treatment of cancer and
 XX helminthic infections.
 XX Sequence 15 AA;
 SQ Query Match 50.0%; Score 74; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITEL 15
 DB 1 QYIKANSKFIGITEL 15

```
Db      1 QYIKANSKFIGITEL 15

RESULT 11
AAV82637
ID      AAY82637 standard; peptide; 15 AA.
XX
AC      AAY82637;
XX
DT      07-AUG-2000 (first entry)
XX
DE      Tetanus toxoid T cell epitope peptide SEQ ID NO:13.
XX
KW      T cell epitope; B cell epitope; allergy; allergen; antigenic;
KW      anti-allergic; antiasthmatic; antiinflammatory; dermatological;
KW      immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW      atopic dermatitis; acute urticaria; chronic urticaria;
KW      gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW      anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
OS      Clostridium tetani.
XX
PN      Synthetic.
XX
PN      WO200006694-A2.
XX
PD      10-FEB-2000.
XX
PF      20-JUL-1999; 99WO-BE00092.
XX
PR      30-JUL-1998; 98EP-0870167.
XX
PA      (UNTO ) UCB SA.
XX
PI      Saint-Remy J, Jacquemin M;
XX
DR      WPI: 2000-422470/36.
XX
PT      New compound for prevention and treatment of allergies comprises at
PT      least one allergen antigenic determinant recognized by a B cell and at
PT      least one antigenic determinant which does not trigger T cell
PT      activation .
XX
PS      Example 6; Page 30; 50pp; English.
XX
CC      The present invention describes a compound (I) for the prevention and/or
CC      treatment of allergy. The compound comprises at least one allergen
CC      antigenic determinant (I) recognised by a B cell or an antibody secreted
CC      by a B cell of a non-atopic individual and at least one antigenic
CC      determinant (II) different from the allergen that triggers T cell
CC      activation. (I) has anti-allergic, antiasthmatic, antiinflammatory,
CC      dermatological and immunosuppressive activities, and can be used in a
CC      vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
CC      treat and/or prevent allergies or a disease of allergic origin,
CC      especially hypersensitivities. These include rhinitis, sinusitis,
CC      bronchial asthma, atopic dermatitis, some forms of acute and chronic
CC      urticaria, gastro-intestinal syndromes associated with the ingestion of
CC      food allergens, oro-pharyngeal syndrome, anaphylactic reactions
CC      associated with drug hypersensitivities and/or a mixture of these. The
CC      use of (I) in the treatment of allergic conditions avoids the need for
CC      drug treatment, which often causes undesirable side-effects. Also, prior
CC      art drug therapies alleviate symptoms, but do not influence their
CC      causes, however (I) actually combats the cause of an allergic reaction.
CC      The present sequence represents a peptide, which is used in an
CC      example from the present invention.
XX
Sequence 15 AA;
Qy      1 QYIKANSKFIGITEL 15
Db      1 QYIKANSKFIGITEL 15
|||
Query Match 50.0%; Score 74; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
AAV84427
ID      AAY84427 standard; Peptide; 15 AA.
XX
Qy      1 QYIKANSKFIGITEL 15
Db      1 QYIKANSKFIGITEL 15
|||
Query Match 50.0%; Score 74; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      1 QYIKANSKFIGITEL 15

RESULT 12
AAV92625
ID      AAY92625 standard; Protein; 15 AA.
XX
AC      AAY92625;
XX
DT      10-AUG-2000 (first entry)
XX
DE      Foreign epitope P2.
XX
KW      Foreign epitope; P2; prostate specific membrane antigen; PSM; Her2;
KW      Heregulin 2; Fibroblast growth factor 8b; FGF8b; vaccination;
KW      cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW      prostate cancer; cell-associated peptide antigen.
XX
OS      Clostridium tetani.
XX
PN      WO200020027-A2.
XX
PD      13-APR-2000.
XX
PF      05-OCT-1999; 99WO-DK00525.
XX
PR      05-OCT-1998; 98DK-0001261.
XX
PR      20-OCT-1998; 98US-0105011.
XX
PA      (MEBI-) M & E BIOTECH AS.
XX
PI      Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI      Gautam A, Birk P, Karlsson G;
XX
DR      WPI: 2000-349917/30.
XX
DR      N-PSDB: AAA03460.
XX
PT      Inducing immune responses to weakly immunogenic, tumor associated
PT      peptide antigens for the treatment of breast and prostate cancer
XX
PS      Example 1; Page 213; 220pp; English.
XX
CC      The claims detail a method for inducing immune responses against weakly
CC      immunogenic cell-associated peptide antigens (PA) such as those
CC      associated with cancers (i.e. self-proteins), for example, human
CC      prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC      fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC      simultaneous presentation by antigen producing cells (APCs) of the
CC      animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC      group derived from the PA and/or at least 1 B-cell group derived from the
CC      cell-associated PA; and (2) at least 1 first T helper cell group which is
CC      foreign to the animal. Analogues of human PSM, human Her2 and
CC      human/murine FGF8b comprising a substantial part of all known and
CC      predicted CTL and B-cell epitopes of the respective PA and including at
CC      least one foreign T helper epitope (e.g. P2 and/or P30) are also claimed.
CC      The method is used to treat prostate, prostate/breast or breast cancer
CC      when the PA is human PSM, FGF8b and Her2, respectively.
XX
Sequence 15 AA;
Qy      1 QYIKANSKFIGITEL 15
Db      1 QYIKANSKFIGITEL 15
|||
Query Match 50.0%; Score 74; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
AAV84427
ID      AAY84427 standard; Peptide; 15 AA.
XX
```

AA34427;
 25-JUL-2000 (first entry)
 Amino acid sequence of the tetanus toxoid P2 epitope.
 Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;
 tumour necrosis factor receptor; type II transmembrane protein;
 osteoclast differentiation; CSF-1; osteoclast activator;
 immune response; osteoporosis; bone resorption;
 tetanus toxoid P2 epitope.
 Clostridium tetani.
 WO200015807-A1.
 23-MAR-2000.
 13-SEP-1999; 99WO-DK00481.
 15-SEP-1998; 98DK-0001164.
 02-OCT-1998; 98US-0102896.
 (MEBI-) M & E BIOTECH AS.
 Halkier T, Haaning J;
 WPI; 2000-271444/23.
 In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used
 to treat, prevent and ameliorate osteoporosis -
 Example; Page 106; 110pp; English.
 The present sequence represents the tetanus toxoid P2 epitope. It is
 used to create a fusion protein with murine osteoprotegerin ligand
 (OPGL). Osteoprotegerin is a secreted member of the tumour necrosis
 factor receptor family, which blocks osteoclastogenesis in a dose
 dependent manner. The OPGL protein is synthesised as a type II
 transmembrane protein. The murine and human OPGL polypeptides are 87%
 homologous. OPGL is a potent osteoclast differentiation factor when
 combined with CSF-1. It is not capable of inducing osteoclast
 differentiation in the absence of CSF-1. OPGL is also an activator of
 mature osteoclasts. The specification describes a method for the in vivo
 down-regulation of OPGL activity in an animal. The method comprises
 using at least one OPGL polypeptide or subsequence, and/or at least one
 OPGL analogue to induce an immune response in the animal. The method
 and OPGL polypeptide are useful for treating, preventing and ameliorating
 osteoporosis or other diseases or conditions characterised by excessive
 bone resorption.
 Sequence 15 AA;
 Query Match 50.0%; Score 74; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITEL 15
 Db 1 QYIKANSKFIGITEL 15
 RESULT 14
 AA70300
 ID AA70300 standard; peptide; 15 AA.
 AC AA70300;
 XX
 XX
 06-JUN-2000 (first entry)
 Clostridium tetani tetanus toxoid T-cell epitope, P589.
 Recombinant protein; CDC/NIMNIVAC-1; multivalent; malaria; vaccine;
 XX

T-cell epitope; tetanus toxoid; antigenic epitope; treatment;
 circumsporozoite protein; CSP; sporozoite surface protein-2; SSP-2;
 liver stage antigen-1; LSA-1; merozoite surface protein-1; MSP-1; MSP-2;
 apical membrane antigen-1; AMA-1; erythrocyte binding antigen-175;
 EBA-175; rhoptry associated protein-1; RAP-1; gamete specific antigen;
 Pf27; antiparasitic; prevention; anti-CDC/NIMNIVAC-1 antibody.
 Clostridium tetani.
 OS
 WO200011179-A1.
 02-MAR-2000.
 19-AUG-1999; 99WO-US18869.
 21-AUG-1998; 98US-0097703.
 (NAIM-) NAT INST IMMUNOLOGY.
 (USSH) US DEPT HEALTH & HUMAN SERVICES.
 Lal AA, Shi YP, Hasnain SE;
 WPI; 2000-237654/20.
 Novel recombinant protein as vaccine for treating malarial infection
 comprises antigenic peptides obtained from different stages of
 plasmodium falciparum life cycle -
 Claim 2; Page 17; 52pp; English.
 The present sequence is the tetanus toxoid T-cell epitope P589, derived
 from Clostridium tetani. It is used in the construction of recombinant
 protein CDC/NIMNIVAC-1, which is a multivalent, multistage malarial
 vaccine. The recombinant protein comprises, melittin signal peptide,
 (His)6 tag, T-cell epitope from tetanus toxoid and 21 antigenic epitopes
 from circumsporozoite protein (CSP), sporozoite surface protein-2
 (SSP-2), liver stage antigen-1 (LSA-1), merozoite surface protein-1
 (MSP-1), MSP-2, apical membrane antigen-1 (AMA-1), erythrocyte binding
 antigen-175 (EBA-175), rhoptry associated protein-1 (RAP-1) and gamete
 specific antigen, Pf27. These epitopes were obtained at different stages
 of the life cycle of P. falciparum. CDC/NIMNIVAC-1 vaccine has
 antiparasitic activity and can be used for treatment and prevention of
 malarial infections. Anti-CDC/NIMNIVAC-1 antibodies can be used for
 detecting P. falciparum in biological samples.
 Sequence 15 AA;
 Query Match 50.0%; Score 74; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITEL 15
 Db 1 QYIKANSKFIGITEL 15
 RESULT 15
 AA44763
 ID AA44763 standard; peptide; 15 AA.
 XX
 AC AA44763;
 XX
 04-MAY-2000 (first entry)
 Tetanus toxoid protein derived T-cell activating epitope P2.
 Surface layer protein; S-layer secretion signal; antibiotic; vaccine;
 recombinant fusion protein cleavage; enzyme; protein polymer; foodstuff;
 antibacterial enzyme; surface glycoprotein; T-cell activating epitope;
 P2; tetanus toxoid; IPNV; Infectious pancreatic necrosis virus.
 Clostridium tetani.
 OS
 XX

PN WO200004170-A1.
 XX 27-JAN-2000.
 XX 14-JUL-1999; 99WO-CA00637.
 XX 14-JUL-1998; 98CA-2237704.
 XX (UYBR-) UNIV BRITISH COLUMBIA.
 XX Smit J;
 XX WPI; 2000-182434/16.
 XX
 XX Cleavage of Caulobacter produced recombinant fusion proteins useful for
 XX producing vaccine peptides
 XX
 XX Example 2; Page 16; 33pp; English.
 XX
 XX The patent discloses a method for cleaving a recombinant fusion protein
 XX which is produced by Caulobacter and consists of Caulobacter surface
 XX layer (S-layer) protein (containing the C-terminal secretion signal) and
 XX a target protein heterologous to Caulobacter. The cleavage of target
 XX protein from the S-layer protein is carried out under mild acid
 XX conditions so that cleavage occurs at aspartate-proline dipeptide site
 XX without solubilising the protein. The cleavage is accomplished while the
 XX fusion protein is in an insoluble aggregate form which facilitates
 XX purification of the protein. The method is useful for producing pure
 XX proteins including recombinant human and animal therapeutic antibiotic
 XX and vaccine peptides, enzymes, protein polymers, and antibacterial
 XX enzymes for foodstuffs.
 XX The present sequence is a T-cell activating epitope P2 derived
 XX from tetanus toxoid protein. This sequence was fused to a DNA encoding
 XX a fragment of Infectious pancreatic necrosis virus surface glycoprotein
 XX which is a vaccine candidate. This chimeric protein was in turn fused to
 XX DNA encoding C. crescentus S-layer secretion signal (corresponds to the
 XX C-terminal portion of the S-layer protein from amino acid 690 onwards and
 XX contains native Asp-Pro site) for construction of a recombinant
 XX fusion construct which is expressed in Caulobacter and then cleaved
 XX to recover the vaccine candidate protein.
 XX
 XX Sequence 15 AA;
 XX
 XX Query Match 50.0%; Score 74; DB 21; Length 15;
 XX Best Local Similarity 100.0%; Pred. No. 0.0001;
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITEL 15
 Db 1 QYIKANSKFIGITEL 15
 Search completed: October 10, 2002, 17:09:07
 Job time : 27 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:08:27 ; Search time 16 Seconds
(without alignments)
42.745 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYKANSKFIGITELGSLHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 140943

Minimum DB seq length: 0

Maximum DB seq length: 28

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112.5	76.0	27	1	US-08-446-692-13
2	112.5	76.0	27	2	US-08-488-351A-13
3	77	52.0	24	5	PCT-US92-07218-25
4	77	52.0	24	5	PCT-US95-02121-110
5	77	52.0	27	5	PCT-US92-07218-32
6	74	50.0	15	2	US-08-319-704-10
7	74	50.0	15	2	US-08-661-052-6
8	74	50.0	15	2	US-08-480-502-7
9	74	50.0	15	4	US-09-046-373-2
10	74	50.0	15	4	US-09-188-082-6
11	74	50.0	15	5	PCT-US93-11703-69
12	74	50.0	16	4	US-09-248-588-55
13	74	50.0	17	1	US-08-446-692-4
14	74	50.0	17	2	US-08-488-351A-4
15	74	50.0	17	3	US-09-100-409A-40
16	74	50.0	17	5	PCT-US95-08596-23
17	74	50.0	17	5	PCT-US95-13841-7
18	70	47.3	14	1	US-08-186-266-5
19	70	47.3	14	1	US-08-305-871A-5
20	70	47.3	14	1	US-08-485-167A-18
21	70	47.3	14	2	US-08-817-933A-9
22	70	47.3	14	5	PCT-US92-07218-15
23	70	47.3	14	5	PCT-US92-07218-30
24	70	47.3	14	5	PCT-US95-02121-95
25	70	47.3	24	5	PCT-US92-07218-31
26	70	47.3	27	5	PCT-US92-07218-26
27	70	47.3	27	5	PCT-US92-07218-27

28	70	47.3	27	5	PCT-US92-07218-28
29	70	47.3	27	5	PCT-US95-02121-111
30	70	47.3	27	5	PCT-US95-02121-112
31	69	46.6	15	2	US-08-661-052-9
32	69	46.6	15	4	US-09-188-082-9
33	68	45.9	19	1	US-08-787-547-41
34	66	44.6	14	4	US-09-082-279B-510
35	66	44.6	14	4	US-09-315-304B-510
36	65	43.9	13	1	US-08-787-547-42
37	60	40.5	13	5	PCT-US94-10257A-26
38	59	39.9	20	4	US-09-046-276-26
39	59	39.9	20	4	US-09-026-276-30
40	58	39.2	20	1	US-07-690-983D-40
41	58	39.2	20	4	US-09-026-276-29
42	58	39.2	20	4	US-09-026-276-31
43	58	39.2	24	1	US-07-690-983D-43
44	57.5	38.9	28	3	US-09-100-414B-74
45	57.5	38.9	28	4	US-09-303-323-74

ALIGNMENTS

RESULT 1
US-08-446-692-13
; Sequence 13, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Xi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/446,692
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-13

Query Match 76.0%; Score 112.5; DB 1; Length 27;
Best Local Similarity 85.2%; Pred. No. 3.3e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY 1 QYKANSKFIGITELGSLHWSYGLRP 27
DB 3 QYKANSKFIGITELE---HWSYGLRP 26

```
RESULT 2
US-08-488-351A-13
; Sequence 13, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-13

Query Match 76.0%; Score 112.5; DB 2; Length 27;
Best Local Similarity 85.2%; Pred. No. 3.3e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 QYIKANSKEIGTELGPSTLHWSYGLRP 27
Db 3 QYIKANSKEIGTELE---HWSYGLRP 26

RESULT 3
PCT-US92-07218-25
; Sequence 25, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitellio, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/07218
; FILING DATE: 19920826
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14137-26-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US92-07218-25

Query Match 52.0%; Score 77; DB 5; Length 24;
Best Local Similarity 88.9%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKEIGTELGPSTL 18
Db 1 QYIKANSKEIGTEPLPS 18

RESULT 4
PCT-US95-02121-110
; Sequence 110, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
```

;; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-02121-110

Query Match 52.0%; Score 77; DB 5; Length 24;
Best Local Similarity 88.9%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGPS 18
|||||
DB 1 QYIKANSKFIGITEFLPS 18

RESULT 5

PCT-US92-07218-32
; Sequence 32, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitellio, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/07218
; FILING DATE: 19920826
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14137-26-3
; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 415-326-2400
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-07218-32

Query Match 52.0%; Score 77; DB 5; Length 27;
Best Local Similarity 88.9%; Pred. No. 1.3e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGPS 18
|||||
DB 4 QYIKANSKFIGITEFLPS 21

RESULT 6

US-08-319-704-10
; Sequence 10, Application US/08319704
; Patent No. 5814617
; GENERAL INFORMATION:
; APPLICANT: Hoffman, Stephen L.
; APPLICANT: Charoenvit, Yupin
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Doolan, Denise L.
; TITLE OF INVENTION: Protective 17 kDa Malaria Hepatic and
; TITLE OF INVENTION: Erythrocytic Stage Immunogen and Gene
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Naval Medical R & D Command
; STREET: Bldg 1, T-12, 8901 Wisconsin Avenue
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: U.S.A
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/319,704
; FILING DATE: 07-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: A. David Spevack
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: 75,206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 295-6759
; TELEFAX: (301) 295-1022
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-319-704-10

Query Match 50.0%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
|||||
DB 1 QYIKANSKFIGITEL 15

RESULT 7

US-08-661-052-6
; Sequence 6, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: Internal
US-08-661-052-6

Query Match 50.0%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 8
US-08-460-502-7
; Sequence 7, Application US/08460502
; Patent No. 5843464
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Golrick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-502-7

Query Match 50.0%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 9
US-09-046-373-2
; Sequence 2, Application US/09046373
; Patent No. 6235714
; GENERAL INFORMATION:
; APPLICANT: Sudhir Paul
; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Gololobov
; TITLE OF INVENTION: Methods for Identifying Inducers and
; TITLE OF INVENTION: Inhibitors of Catalytic Antibodies, Compositions and Their
; TITLE OF INVENTION: Use
; FILE REFERENCE: UNMC 63123
; CURRENT APPLICATION NUMBER: US/09/046,373
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-046-373-2

Query Match 50.0%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 10
US-09-188-082-6
; Sequence 6, Application US/09188082
; Patent No. 6270765
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES

; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,082
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/661,052
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-188-082-6

Query Match 50.0%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 11
PCT-US93-11703-69
; Sequence 69, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259

; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-11703-69

Query Match 50.0%; Score 74; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 12
US-09-248-588-55
; Sequence 55, Application US/09248588
; Patent No. 6231864
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and
; FILE REFERENCE: SYN-101 4564/69529
; CURRENT APPLICATION NUMBER: US/09/248,588
; CURRENT FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: 60/074537
; EARLIER FILING DATE: 1998-02-12
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Clostridium tetani
; PUBLICATION INFORMATION:
; JOURNAL: Vaccine
; VOLUME: 15
; ISSUE: 4
; PAGES: 377-
; DATE: 1997
US-09-248-588-55

Query Match 50.0%; Score 74; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 13
US-08-446-692-4
; Sequence 4, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York

STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-4

Query Match 50.0%; Score 74; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 3 QYIKANSKFIGITEL 17

RESULT 14
US-08-488-351A-4
Sequence 4, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166

FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-4

Query Match 50.0%; Score 74; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 3 QYIKANSKFIGITEL 17

RESULT 15
US-09-100-409A-40
Sequence 40, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
TITLE OF INVENTION: IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-40

Query Match 50.0%; Score 74; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15

Db 3 QYIRANSKFIGITEL 17
|||||

Search completed: October 10, 2002, 17:11:44
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:12:48 ; Search time 16 Seconds
(without alignments)
198.184 Million cell updates/sec

Title: US-09-848-834A-12

Perfect score: 169

Sequence: 1 DEKKIAKMEKASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 7172

Minimum DB seq length: 0

Maximum DB seq length: 33

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	30.8	10	1 RHPGG	gonadoliberin - pi
2	52	30.8	10	1 RHSHG	gonadoliberin - sh
3	48	28.4	10	1 RHAQ1	gonadoliberin I -
4	39	23.1	10	2 A21114	gonadoliberin - ch
5	36	21.3	30	2 F70254	hypothetical prote
6	34	20.1	10	1 RHAQ2	gonadoliberin II -
7	34	20.1	10	1 A61126	gonadoliberin - sp
8	34	20.1	10	2 A46030	gonadoliberin I -
9	34	20.1	10	2 B46030	gonadoliberin II -
10	34	20.1	22	2 F44957	osmotin homolog -
11	34	20.1	22	2 B33174	sorhamin - sorghum
12	33.5	19.8	32	2 S20771	Ig heavy chain v r
13	32	18.9	16	2 S57517	T cell receptor be
14	31	18.3	10	2 A49187	gonadotropin-relea
15	30	17.8	21	2 C60275	fibrillin - porph
16	30	17.8	33	2 A56818	Na+/K+-exchanging
17	29	17.2	21	2 D60275	fibrillin - porph
18	29	17.2	30	2 A49955	protein-tyrosine k
19	28	16.6	22	2 P00070	T-cell receptor va
20	28	16.6	28	2 A49829	T-cell receptor va
21	28	16.6	31	2 G82816	hypothetical prote
22	28	16.6	32	2 B70241	hypothetical prote
23	27.5	16.3	30	2 PC2340	second envelope ql
24	27.5	16.3	30	2 S15650	NADH dehydrogenase
25	27	16.0	17	2 I51203	myosin heavy chain
26	27	16.0	18	2 A40256	interleukin-7 rece
27	27	16.0	22	2 B40256	interleukin-7 rece
28	27	16.0	25	2 B36934	orf3 3' of madA -
29	27	16.0	27	2 C69350	hypothetical prote

30	27	16.0	31	2 S32610	antiviral protein
31	27	16.0	31	2 H95177	hypothetical prote
32	27	16.0	32	2 A60363	apolipoprotein III
33	27	16.0	32	4 C28671	hypothetical chl p
34	27	16.0	33	2 S05404	alpha-2-macroglobu
35	27	16.0	33	2 G95006	hypothetical prote
36	26.5	15.7	27	2 A25511	Cc gene smaller hy
37	26	15.4	20	2 S00315	photosystem I chai
38	26	15.4	20	2 PC1151	equinotoxin IC - s
39	26	15.4	20	2 B34016	tenebrosin B - sea
40	26	15.4	21	2 S28436	major outer membra
41	26	15.4	22	2 PH1325	Ig heavy chain DJ
42	26	15.4	25	2 PQ0683	photosystem I 16.1
43	26	15.4	28	2 B56779	tetM 5'-region lea
44	26	15.4	30	2 S74121	fructose-bisphosph
45	26	15.4	30	2 S29138	aniline monooxygen

ALIGNMENTS

RESULT 1

RHPGG

gonadoliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase m

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Note: the synthetic and natural hormones have the same physicochemical and biologic

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: trp-3 appears to be essential for biological activity

C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 30.8%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.26;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32

DB 2 HWSYGLRP 9

RESULT 2

RHSHG

gonadoliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

A;Residues: 1-10 <BUR>
A;Note: the natural and synthetic hormones have the same biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyroglutamic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 30.8%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
|||||||
Db 2 HWSYGLRP 9

RESULT 3
RHAQ1
gonadoliberin I - American alligator
N;Alternate names: gonadotropin-releasing hormone I
C;Species: Alligator mississippiensis (American alligator)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C;Accession: A60066
R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Regul. Pept. 33, 105-116, 1991
A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of the American alligator
A;Reference number: A60066; MUID:91352338
A;Accession: A60066
A;Molecule type: protein
A;Residues: 1-10 <LOV>
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyroglutamic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.4%; Score 48; DB 1; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.98;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
|||||||
Db 2 HWSYGLRP 9

RESULT 4
RHAQ2
gonadoliberin - chum salmon
C;Species: Oncorhynchus keta (chum salmon)
C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C;Accession: A21114
R;Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A;Title: Characterization of a teleost gonadotropin-releasing hormone.
A;Reference number: A21114; MUID:83195140
A;Accession: A21114
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <SHE>

Query Match 23.1%; Score 39; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
|||||||
Db 2 HWSYGLRP 9

RESULT 5
F70254
hypothetical protein BBK29 - Lyme disease spirochete plasmid K/lp36
C;Species: Borrelia burgdorferi (Lyme disease spirochete)

A;Residues: 1-10 <BUR>
A;Note: the natural and synthetic hormones have the same biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyroglutamic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 30.8%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
|||||||
Db 2 HWSYGLRP 9

RESULT 3
RHAQ1
gonadoliberin I - American alligator
N;Alternate names: gonadotropin-releasing hormone I
C;Species: Alligator mississippiensis (American alligator)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C;Accession: A60066
R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Regul. Pept. 33, 105-116, 1991
A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of the American alligator
A;Reference number: A60066; MUID:91352338
A;Accession: A60066
A;Molecule type: protein
A;Residues: 1-10 <LOV>
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyroglutamic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.4%; Score 48; DB 1; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.98;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
|||||||
Db 2 HWSYGLRP 9

RESULT 4
RHAQ2
gonadoliberin - chum salmon
C;Species: Oncorhynchus keta (chum salmon)
C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C;Accession: A21114
R;Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A;Title: Characterization of a teleost gonadotropin-releasing hormone.
A;Reference number: A21114; MUID:83195140
A;Accession: A21114
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <SHE>

Query Match 23.1%; Score 39; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
|||||||
Db 2 HWSYGLRP 9

RESULT 5
F70254
hypothetical protein BBK29 - Lyme disease spirochete plasmid K/lp36
C;Species: Borrelia burgdorferi (Lyme disease spirochete)

C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C;Accession: F70254
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wilson, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Venter, A.; Smith, H.O.; Venter, J.C.
Nature 390, 580-586, 1997
A;Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943
A;Accession: F70254
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-30 <KLE>
A;Cross-references: GB:AE000788; NID:g2690123; PIDN:AAC66163.1; PID:g2690155; TIGR:BE000000
A;Experimental source: strain B31
C;Genetics:
A;Genome: plasmid

Query Match 21.3%; Score 36; DB 2; Length 30;
Best Local Similarity 35.3%; Pred. No. 1.7e+02;
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 DEKKIAKMEKASSVFNV 17
|||||
Db 9 DKEKMKIRKLSSYYKI 25

RESULT 6
RHAQ2
gonadoliberin II - American alligator
N;Alternate names: gonadotropin-releasing hormone II
C;Species: Alligator mississippiensis (American alligator)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C;Accession: B60066
R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Regul. Pept. 33, 105-116, 1991
A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of the American alligator
A;Reference number: A60066; MUID:91352338
A;Accession: B60066
A;Molecule type: protein
A;Residues: 1-10 <LOV>
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyroglutamic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 20.1%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
|||||
Db 2 HWSYGLRP 9

RESULT 7
RHAQ2
gonadoliberin - spotted ratfish
N;Alternate names: gonadotropin-releasing hormone
C;Species: Hydrolagus colliet (spotted ratfish)
C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C;Accession: A61126
R;Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.
Gen. Comp. Endocrinol. 82, 152-161, 1991
A;Title: Primary structure of gonadotropin-releasing hormone from the brain of a holbrookei
A;Reference number: A61126; MUID:91340067
A;Accession: A61126
A;Molecule type: protein
A;Residues: 1-10 <LOV>
A;Experimental source: brain
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F;1/Modified site: pyroglutamic acid (Gln) #status experimental

F.10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 20.1%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32

Db 2 HWSHGWP 9

RESULT 8

A46030

gonadoliberin I - spiny dogfish

N:Alternate names: gonadotropin-releasing hormone

C:Species: Squalus acanthias (spiny dogfish)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998

C:Accession: A46030

R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R

Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992

A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro

A:Reference number: A46030; MUID:92335300

A:Accession: A46030

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <LOV>

C:Keywords: hormone; pyroglutamic acid

F.1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 20.1%; Score 34; DB 2; Length 10;

Best Local Similarity 62.5%; Pred. No. 98;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32

Db 2 HWSHGWP 9

RESULT 9

B46030

gonadoliberin II - spiny dogfish

N:Alternate names: gonadotropin-releasing hormone

C:Species: Squalus acanthias (spiny dogfish)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998

C:Accession: B46030

R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R

Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992

A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro

A:Reference number: A46030; MUID:92335300

A:Accession: B46030

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <LOV>

C:Keywords: hormone; pyroglutamic acid

F.1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 20.1%; Score 34; DB 2; Length 10;

Best Local Similarity 62.5%; Pred. No. 98;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32

Db 2 HWSHGWP 9

RESULT 10

F44957

osmotin homolog - common tobacco (cv. Samsun NN) (fragment)

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 12-Apr-1995

C:Accession: F44957

R:Takeda, S.; Sato, F.; Ida, K.; Yamada, Y.

Plant Cell Physiol. 31, 215-221, 1990

A:Title: Characterization of polypeptides that accumulate in cultured Nicotiana tabac
A:Reference number: A44957
A:Accession: F44957
A>Status: Preliminary
A:Molecule type: protein
A:Residues: 1-22 <TAK>
C:Superfamily: thaumatin I

Query Match 20.1%; Score 34; DB 2; Length 22;

Best Local Similarity 43.8%; Pred. No. 2.4e+02;

Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 12 SSFVNVVNSGPSLHWS 27

Db 1 SGVFEVHNXPYTVWA 16

RESULT 11

B31174

sorhamin - sorghum (fragment)

C:Species: Sorghum bicolor (sorghum)

C:Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 12-Apr-1995

C:Accession: B31174

R:Vigers, A.J.; Roberts, W.K.; Sellitrennikoff, C.P.

submitted to the Protein Sequence Database, May 1991

A:Reference number: A33174

A:Accession: B31174

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-22 <VIG>

C:Superfamily: thaumatin I

Query Match 20.1%; Score 34; DB 2; Length 22;

Best Local Similarity 46.7%; Pred. No. 2.4e+02;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 13 SVFNVVNSGPSLHWS 27

Db 1 AVFTVVNRCPYTVWA 15

RESULT 12

S20771

Ig heavy chain V region (VH4, IN1P24) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 23-Jul-1999

C:Accession: S20771

R:Mortari, F.; Wang, J.; Schroeder, H.W.

submitted to the EMBL Data Library, April 1992

A:Description: Analysis of human cord blood Ig heavy chain IgA and IgG repertoire.

A:Reference number: S20764

A:Accession: S20771

A:Molecule type: DNA

A:Residues: 1-32 <MOR>

C:Cross-references: EMBL:Z11955; NID:g33871; PIDN:CAA78012.1; PID:g33872

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 19.8%; Score 33.5; DB 2; Length 32;

Best Local Similarity 70.0%; Pred. No. 4.3e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 20 SGPSLHWSYG 29

Db 15 SGPSL-WDWG 23

RESULT 13

S57517

T cell receptor beta chain V region - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999

C:Accession: S57517

R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Arguet, V.P.
submitted to the EMBL Data Library, June 1995

A:Description: T cell receptor repertoire for a viral epitope in humans is diversified b

A:Reference number: S57494

A:Accession: S57517

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-16 <BUR>

A:Cross-references: EMBL:49924; NID:g887480; PIDN:CAA90170.1; PID:g887481

C:Keywords: T-cell receptor

Query Match 18.9%; Score 32; DB 2; Length 16;

Best Local Similarity 60.0%; Pred. No. 3.2e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 20 SGPSPHWSYG 29

|||||

Db 3 GPPPLRGNYG 12

RESULT 14

A49187

gonadotropin-releasing hormone III - sea lamprey

C:Species: Petromyzon marinus (sea lamprey)

C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995

C:Accession: A49187

R;Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.

Endocrinology 132, 1125-1131, 1993

A:Title: Primary structure and biological activity of a third gonadotropin-releasing hor

A:Reference number: A49187; MUID:93178316

A:Accession: A49187

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <SOW>

A:Experimental source: brain

A:Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 18.3%; Score 31; DB 2; Length 10;

Best Local Similarity 50.0%; Pred. No. 2.6e+02;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32

|||||

Db 2 HWSHDWKP 9

RESULT 15

C60275

fimbrillin - Porphyromonas gingivalis (strain FAY 19W-1) (fragment)

C:Species: Porphyromonas gingivalis

C:Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 11-Jan-2000

C:Accession: C60275

R;Lee, J.Y.; Sojar, H.T.; Bedi, G.S.; Genco, R.J.

Infect. Immun. 59, 383-389, 1991

A:Title: Porphyromonas (bacteroides) gingivalis fimbrillin: size, amino-terminal sequen

A:Reference number: A60275; MUID:91099990

A:Accession: C60275

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-21 <LEE>

C:Superfamily: Porphyromonas fimbrillin

Query Match 17.8%; Score 30; DB 2; Length 21;

Best Local Similarity 62.5%; Pred. No. 8.4e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEKTIAM 8

|||||

Db 7 DESKVAKL 14

Search completed: October 10, 2002, 17:17:19

Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:11:12 ; Search time 13 Seconds
(without alignments)
98.288 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKKIAKMEKASSVFNVNVSGPSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 2152

Minimum DB seq length: 0

Maximum DB seq length: 33

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	48	28.4	10	GONL_ALIMI	P37041 alligator m
2	41	24.3	10	GONL_CLUPA	P81749 clupea pall
3	39	23.1	10	GON3_ONCKE	P20367 oncorhynch
4	36	21.3	29	TLP_ACTDE	P81370 actinidia d
5	34	20.1	10	GON2_CHICK	P37043 gallus gall
6	34	20.1	10	GONL_SQUAC	P27429 squalus aca
7	31	18.3	10	GON3_PETMA	P30948 petromyzon
8	30	17.8	21	SPI3_SOLTU	P58516 solanum tub
9	29	17.2	33	YCI2_NEPOL	Q9tky6 nephroselm
10	27	16.0	18	SODM_MYCHA	P80582 mycobacteri
11	27	16.0	27	Y803_ARCFU	O29455 archaeoglob
12	26	15.4	10	GONL_CHEPR	P80677 chelyosoma
13	26	15.4	16	FOR1_WYRGU	P81438 myrmecia gu
14	26	15.4	20	PSAF_PEA	P20119 pisum sativ
15	26	15.4	20	TENB_ACTTE	P30834 actinia ten
16	26	15.4	26	PGW4_PACGO	P82426 pachycondyl
17	26	15.4	32	APL3_DIAGR	P81471 diatraea gr
18	25.5	15.1	23	COXK_CANFA	Q9tr28 canis fami
19	25.5	15.1	24	KPK_CLOPA	P81344 clostridium
20	25	14.8	8	ALL1_CYDPO	P82152 cydia pomon
21	25	14.8	10	GRP_RANRI	P23260 rana ridibu
22	25	14.8	12	UR2A_CATCO	P04558 catostomus
23	25	14.8	13	SODM_CANFA	P54712 canis fami
24	25	14.8	20	JHBP_BOMMO	P81627 bombyx mori
25	25	14.8	23	GRP_ONCMY	Q9ps30 oncorhynch
26	25	14.8	24	SODC_RANCA	P23417 rana catesb
27	25	14.8	25	GRP_SCYCA	P09472 scyliorhinu
28	25	14.8	27	GRP_CANFA	P08989 canis fami
29	25	14.8	27	GRP_CHICK	P01295 gallus gall
30	25	14.8	27	GRP_PIG	P01294 sus scrofa
31	25	14.8	28	GRP_ALIMI	P31886 alligator m
32	25	14.8	28	PA22_MICNI	P21791 micrurus ni
33	25	14.8	28	VI03_VACCP	Q00334 vaccinia vi

ALIGNMENTS

RESULT 1

GONL_ALIMI
ID GONL_ALIMI STANDARD; PRT; 10 AA.
AC P37041: P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RA MEDLINE=91352338; PubMed=1882082;
RX Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
Regul. Pept. 33:105-116(1991).
RL
CC - FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC - SUBCELLULAR LOCATION: Secreted.
CC - SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Hormone; Amigaion; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match

Best Local Similarity 28.4%; Score 48; DB 1; Length 10;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32

|||||:1

Db 2 HWSYGLQP 9

RESULT 2

GONL_CLUPA

ID GONL_CLUPA STANDARD; PRT; 10 AA.

AC P81749;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH)

DE (Luliberin I).

GN GNRH1.

OS Clupea pallasi (Pacific herring).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;

```
OC Clupeinae: Clupea.
OX NCBI_TaxID=30724;
RN [1]
RC TISSUE-Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
  Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
  hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH: 1.
DR PROSITE: PS00473; GnRH: 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 AMIDATION.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;

Query Match 24.3%; Score 41; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 5.2;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
   |||:|:|
DB 2 HWSHGLSP 9

RESULT 3
ID GON3_ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GnRH-III) (LH-
  RH III) (Luliberin III).
GN GnRH3.
OS Oncorhynchus keta (Chum salmon), and
OS Clupea pallasii (Pacific herring).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O.keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone."
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
  Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
  hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
  FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC PIR: A21114; A21114.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH: 1.
DR PROSITE: PS00473; GnRH: 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

us-09-848-834a-12.closed.rsp

FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1230 MW; 284B323378B45A3 CRC64;

Query Match 23.1%; Score 39; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 9.9;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
   |||:|:|
DB 2 HWSYGLWP 9

RESULT 4
ID TLP_ACTDE STANDARD; PRT; 29 AA.
AC P81370;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thaumatin-like protein (Fragment).
OS Actinidia deliciosa (Kiwi).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Ericales; Actinidiaceae; Actinidia.
OX NCBI_TaxID=3627;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. HAYWARD; TISSUE=Stem plug;
RC WURMS K.V., Greenwood D.R., Sharrock K.R., Long P.G.;
RT "Thaumatin-like protein in kiwifruit."
RL J. Sci. Food Agric. 79:1448-1452(1999).
CC -!- TISSUE SPECIFICITY: WOODY STEM PLUG.
CC -!- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
DR HSP: P25871; LAUN.
DR InterPro: IPR001938; Thaumatin.
DR Pfam: PF00314; Thaumatin; 1.
DR ProDom: PD001321; Thaumatin; 1.
DR PROSITE: PS00316; THAUMATIN; PARTIAL.
FT NON_TER 29 29
FT SEQUENCE 29 AA; 3074 MW; E99F916F60AC377B CRC64;

Query Match 21.3%; Score 36; DB 1; Length 29;
Best Local Similarity 33.3%; Pred. No. 80;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 13 SVFNVVNSGPSLHWS 27
   :||::|:|
DB 1 ATFNIINCPFTVWA 15

RESULT 5
ID GON2_CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GnRH-II)
  (LH-RH II) (Luliberin II).
OS Gallus gallus (Chicken),
OS Alligator mississippiensis (American alligator),
OS Squalus acanthias (Spiny dogfish),
OS Hydrolaqus collii (Spotted ratfish) (Pacific ratfish), and
OS Clupea pallasii (Pacific herring).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=Chicken; TISSUE=Hypothalamus;
RX MEDLINE=84222059; PubMed=6427779;
RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
```

RA Matsuo H.;
RT "Identification of the second gonadotropin-releasing hormone in
RT chicken hypothalamus: evidence that gonadotropin secretion is
RT probably controlled by two distinct gonadotropin-releasing hormones
RT in avian species.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=A.mississippiensis; TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
RT from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
RN [3]
RP SEQUENCE.
RC SPECIES=S.acanthias; TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
RT dogfish brain provides insight into GnRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN [4]
RP SEQUENCE.
RC SPECIES=H.collietii; TISSUE=Brain;
RX MEDLINE=91340067; PubMed=1678723;
RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
RA Lee T.;
RT "Primary structure of gonadotropin-releasing hormone from the brain
RT of a holoccephalan (ratfish: Hydrolagus colliei).";
RL Gen. Comp. Endocrinol. 82:152-161(1991).
RN [5]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasi; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RT hormones, including a novel form, from an ancient teleost, herring";
RL Endocrinology 141:509-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: B60066; RHAQ2
DR PIR: A61126; A61126.
DR PIR: B46030; B46030.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;
Query Match 20.1%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 25 HWSYGLRP 32
Db 2 HWSHGWP 9
|||||
RESULT 6
GONL_SQUAC STANDARD; PRT; 10 AA.
AC P27429;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin (Gonadotropin-releasing hormone) (GnRH) (LH-RH)

DE (Luliberin).
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
RT dogfish brain provides insight into GnRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;
Query Match 20.1%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 25 HWSYGLRP 32
Db 2 HWSHGWP 9
|||||
RESULT 7
GON3_PETMA STANDARD; PRT; 10 AA.
AC P30948;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GnRH-III)
DE (Luliberin III).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
RT releasing hormone from lamprey brain.";
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237AA1F5A3 CRC64;
Query Match 18.3%; Score 31; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 25 HWSYGLRP 32
|||||

```
Db 2 HWSHMKP 9

RESULT 8
SPI3_SOLTU STANDARD; PRT; 21 AA.
AC P58516;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Serine protease inhibitor 3 (SPPI-22) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Istrinskii;
RX MEDLINE=98258940; PubMed=9598993;
RA Valueva T.A., Revina T.A., Kladnitskaya G.V., Mosolov V.V.;
RT "Kunitz-type proteinase inhibitors from intact and
RT Phytophthora-infected potato tubers.";
RL FEBS Lett. 426:131-134(1998).
CC -|- FUNCTION: INHIBITS TRYPSIN AND CHYMOTRYPSIN (SERINE PROTEASES).
CC DOES NOT INHIBIT ELASTASE, SUBTILISIN, CATHEPSIN L NOR PAPAINE
CC (SERINE AND CYSTEINE PROTEASES). PROTECTS THE PLANT BY INHIBITING
CC PROTEASES OF INVADING ORGANISMS, DECREASING BOTH HYPHAL GROWTH AND
CC ZOOSPORES GERMINATION OF PHYTOPHTHORA INFESTANS.
CC -|- SUBCELLULAR LOCATION: VACUOLAR (BY SIMILARITY).
CC -|- TISSUE SPECIFICITY: TUBERS.
CC -|- INDUCTION: By infection with Phytophthora infestans.
CC -|- MISCELLANEOUS: HAS A SINGLE CHAIN STRUCTURE.
CC -|- SIMILARITY: BELONGS TO THE LEGUMINOUS KUNITZ-TYPE INHIBITOR
CC FAMILY.
DR PROSITE; PS00283; SOYBEAN_KUNITZ; PARTIAL.
KW Serine protease inhibitor; Multigene family.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2439 MW; 5674BEAB7149147D CRC64;

Query Match 17.8%; Score 30; DB 1; Length 21;
Best Local Similarity 43.8%; Pred. No. 4.1e-02;
Matches 7; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 15 FNVVNSGSLH--WSY 28
|:|:| | |
Db 5 FDLVDSGRDLRGWY 20

RESULT 9
YC12_NEPOL STANDARD; PRT; 33 AA.
AC Q9TKV6;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 3.5 kDa protein ycf12 (RF12).
GN ycf12.
OS Nephroselmis olivacea.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
OC Chlorodendrales; Chlorodendraceae; Nephroselmis.
OX NCBI_TaxID=31312;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-484;
RX MEDLINE=99398694; PubMed=10468594;
RA Turmel M., Otis C., Lemieux C.;
RT "The complete chloroplast DNA sequence of the green alga Nephroselmis
RT olivacea: insights into the architecture of ancestral chloroplast
RT genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).
CC -|- SIMILARITY: BELONGS TO THE YCF12 FAMILY.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF137379; RAD54830.1;
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 33 AA; 3412 MW; DBFFB07A8B9F20D3 CRC64;

Query Match 17.2%; Score 29; DB 1; Length 33;
Best Local Similarity 60.0%; Pred. No. 9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 13 SVFNVVNSGP 22
|:| | | |
Db 11 SLFAVVAAGP 20

RESULT 10
SODM_MYCHA STANDARD; PRT; 18 AA.
AC P80582;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn] (EC 1.15.1.1) (fragment).
GN SODA OR SOD.
OS Mycobacterium habana.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1784;
RN [1]
RP SEQUENCE.
RC STRAIN=TMC 5135;
RX MEDLINE=96262709; PubMed=8704977;
RA Bisht D., Mehrotra J., Dhindsa M.S., Singh N.B., Sinha S.;
RT "A major T-cell-inducing cytosolic 23 kDa protein antigen of the
RT vaccine candidate Mycobacterium habana is superoxide dismutase.";
RL Microbiology 142:1375-1383(1996).
CC -|- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -|- COFACTOR: Manganese (By similarity).
CC -|- SUBUNIT: HOMODIMER (PROBABLE).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR HSP; P17670; 1IDS.
DR InterPro; IPR001189; SOD_MI.
DR Pfam; PF00081; sodfe; 1.
DR PROSITE; PS00088; SOD_MN; PARTIAL.
KW Oxidoreductase; Manganese.
FT NON_TER 18
SQ SEQUENCE 18 AA; 1883 MW; A4161A3DAC93F710 CRC64;

Query Match 16.0%; Score 27; DB 1; Length 18;
Best Local Similarity 57.1%; Pred. No. 9.2e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 22 PSLHWSY 28
| | | |
Db 6 PDLGWDY 12

RESULT 11
Y803_ARCFU STANDARD; PRT; 27 AA.
ID Y803_ARCFU
AC Q29455;
DT 16-OCT-2001 (Rel. 40, Created)
```


DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0803.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE01049; AAB90446.1; -
DR TIGR; AF0803; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 27 AA; 3203 MW; 913FF63AB470903A CRC64;

Query Match 16.0%; Score 27; DB 1; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 ASSVFNVVNS 20
I:| | | | |
Db 8 ATSKFNIVKN 17

RESULT 12
GON1_CHEPR
ID GON1_CHEPR STANDARD; PRT; 10 AA.
AC P80677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I)
DE (Luliberin I).
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
RT and the evolutionary implications."
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
CC THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -!- MASS SPECTROMETRY: MW=1246.56; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro; IPR002012; GNRH.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
SQ SEQUENCE 10 AA; 1264 MW; 284B3639DBSAB5A3 CRC64;

Query Match 15.4%; Score 26; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
| | | | |
Db 2 HWSDFPK 9

RESULT 13
FOR1_MYRGU
ID FOR1_MYRGU STANDARD; PRT; 16 AA.
AC P81438;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Formaeicin 1.
OS Myrmecia gulosa (Red bulldog ant).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Formicidae; Myrmecinae; Myrmecia.
OX NCBI_TaxID=36170;
RN [1]
RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
RC TISSUE=Hemolymph;
RX MEDLINE=98165787; PubMed=9497332;
RA Mackintosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
RT "Isolation from an ant Myrmecia gulosa of two inducible
RT O-glycosylated proline-rich antibacterial peptides."
RL J. Biol. Chem. 273:6139-6143(1998).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE. HAS ACTIVITY AGAINST E.COLI
CC BUT NONE AGAINST OTHER GRAM-NEGATIVE BACTERIA AND GRAM-POSITIVE
CC BACTERIA.
CC -!- INDUCTION: By bacterial infection.
CC -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
CC -!- SIMILARITY: TO DROSOPHILA DROSOCIN.
KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
FT CARBOHYD 11 11 O-LINKED (GALNAC...).
SQ SEQUENCE 16 AA; 1794 MW; 80CEA3AABBC2E0AE CRC64;

Query Match 15.4%; Score 26; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 16 NVVNSGPSLH 25
| | | | |
Db 4 NPVNNKPTPH 13

RESULT 14
PSAF_PEA
ID PSAF_PEA STANDARD; PRT; 20 AA.
AC P20119;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Photosystem I reaction centre subunit III (light-harvesting complex I
DE 17 kDa protein) (PSI-F) (Fragment).
GN PSAF.
OS Pisum sativum (Garden pea).

Search completed: October 10, 2002, 17:15:52
Job time : 14 secs

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]

RP SEQUENCE.
RX MEDLINE=88137587; PubMed=3277857;
RA Dunn P.P.J., Packman L.C., Pappin D., Gray J.C.;
RT "N-terminal amino acid sequence analysis of the subunits of pea
RT photosystem I";
RL FEBS Lett. 228:157-161(1988).
CC -!- FUNCTION: PROBABLY PARTICIPATES IN EFFICIENCY OF ELECTRON TRANSFER
CC FROM PLASTOCYANIN TO P700 (OR CYTOCHROME C553 IN ALGAE AND
CC CYANOBACTERIA). THIS PLASTOCYANIN-DOCKING PROTEIN CONTRIBUTES
CC TO THE SPECIFIC ASSOCIATION OF PLASTOCYANIN TO PS I.
CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH LUMENAL SIDE OF THE
CC THYLAKOID MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE PSAF FAMILY.
DR PIR; S00315; S00315.
DR Mendel; 16765; PISsa;PsaF;mm16765.
KW Photosynthesis; Photosystem I; Chloroplast; Thylakoid; Membrane.
FT VARIANT 1 1
FT NON_TER 20 20 A -> D.
SQ SEQUENCE 20 AA; 2250 MW; E2C79E760187898C CRC64;

Query Match 15.4%; Score 26; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DEKKIAKMEK 10
DB 10 ESKQFAKREK 19

RESULT 15

TENB_ACTTE STANDARD; PRT; 20 AA.
AC F30834;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Tenebrosin B (fragment).
OS Actinia tenebrosa (Australian red waratah sea anemone).
OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthaeae; Actiniidae; Actinia.
OX NCBI_TaxID=6105;
RN [1]
RP SEQUENCE.
RX MEDLINE=90232538; PubMed=1970442;
RA Norton R.S., Bobek G., Ivanov J.O., Thomson M., Fiala-Beer E.,
RA Moritz R.L., Simpson R.J.;
RT "Purification and characterisation of proteins with cardiac
RT stimulatory and haemolytic activity from the anemone Actinia
RT tenebrosa";
RL Toxicon 28:29-41(1990).
CC -!- FUNCTION: THIS CARDIAC STIMULATORY AND HEMOLYTIC PROTEIN IS A
CC CHANNEL-FORMING AND/OR MEMBRANE-PENETRATING PROTEIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE TENEBROSIN FAMILY.
DR PIR; B34016; B34016.
KW Hemolysis; Toxin; Transmembrane.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 1960 MW; FA32B426009FF5FA CRC64;

Query Match 15.4%; Score 26; DB 1; Length 20;
Best Local Similarity 46.2%; Pred. No. 1.4e+03;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 AKMEKASSVFNVV 18
DB 7 AVIEGATLTFNVL 19

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:11:53 ; Search time 29 Seconds
(without alignments)
196.856 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKIAKMERASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 15360

Minimum DB seq length: 0
Maximum DB seq length: 33

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	41	24.3	33	13 Q9PT34	Q9pt34 oncorhynchu
2	41	24.3	33	13 Q9W7G0	Q9w7g0 oncorhynchu
3	36	21.3	30	16 O50832	O50832 borrelia bu
4	35	20.7	27	10 Q9S909	Q9s909 zea mays (m
5	34	20.1	22	10 Q9S908	Q9s908 sorghum bic
6	34	20.1	26	10 Q9S8K5	Q9s8k5 nicotiana t
7	33	19.5	25	11 Q9QVA3	Q9qva3 rattus sp.
8	32	18.9	27	12 Q9QHP5	Q9qhp5 hepatitis c
9	32	18.9	33	6 Q95LC4	Q95lc4 sus scrofa
10	31.5	18.6	28	15 Q73626	Q73626 human immun
11	31	18.3	18	2 Q9R4C3	Q9r4c3 agrobacteri
12	31	18.3	24	2 Q56150	Q56150 streptomyce
13	31	18.3	25	4 Q9UD88	Q9ud88 homo sapien
14	31	18.3	33	8 Q9T2T5	Q9t2t5 bos taurus
15	30.5	18.0	27	12 Q9IIP8	Q9ilip8 hepatitis c
16	30	17.8	20	2 O34197	O34197 rickettsia

17	30	17.8	25	4 Q9UC30	Q9uc30 homo sapien
18	30	17.8	25	5 Q9NDS5	Q9nds5 drosophila
19	30	17.8	27	12 Q9QHX3	Q9qh3 hepatitis c
20	30	17.8	28	2 Q9R5C7	Q9r5c7 vibrio para
21	30	17.8	30	2 Q9ETT2	Q9ett2 ureaplasma
22	30	17.8	30	2 Q9ETT1	Q9ett1 ureaplasma
23	30	17.8	32	5 Q95NY7	Q95ny7 drosophila
24	30	17.8	33	13 Q9PS72	Q9ps72 rana catesb
25	29	17.2	21	10 Q9S8K6	Q9s8k6 mesembryant
26	29	17.2	29	15 Q9R348	Q9r348 human immun
27	29	17.2	30	2 Q9S2Z7	Q9s2z7 chlamydia t
28	29	17.2	30	4 Q9UBS6	Q9ubs6 homo sapien
29	29	17.2	30	4 Q9UMI6	Q9umi6 homo sapien
30	29	17.2	32	5 Q9U340	Q9u340 caenorhabd1
31	28.5	16.9	33	5 Q25536	Q25536 notopiana a
32	28	16.6	17	8 Q9T2R9	Q9t2r9 solanum tub
33	28	16.6	17	11 P97758	P97758 mus musculu
34	28	16.6	20	6 P79256	P79256 aotus trivi
35	28	16.6	23	2 Q9R4U4	Q9r4u4 yersinia ps
36	28	16.6	25	4 Q9QUU9	Q9quu9 homo sapien
37	28	16.6	26	12 Q82943	Q82943 polyomaviru
38	28	16.6	27	8 Q9AQ58	Q9aq58 schistosoma
39	28	16.6	27	12 Q9QH11	Q9qh1 hepatitis c
40	28	16.6	27	12 Q9J524	Q9j524 hepatitis c
41	28	16.6	27	12 Q9J522	Q9j522 hepatitis c
42	28	16.6	27	12 Q9J521	Q9j521 hepatitis c
43	28	16.6	27	12 Q9J520	Q9j520 hepatitis c
44	28	16.6	27	12 Q9J519	Q9j519 hepatitis c
45	28	16.6	27	12 Q9J518	Q9j518 hepatitis c

ALIGNMENTS

RESULT 1

Q9PT34 ID Q9PT34 PRELIMINARY: PRT: 33 AA.
AC Q9PT34;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
(LULIBERIN) (FRAGMENT).
GN GNRH1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RT Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110533; AAD3461.1; -.
DR InterPro; IPR002047; AKH.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00256; AKH; UNKNOWN_1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;

```
Query Match 24.3%; Score 41; DB 13; Length 33;
Best Local Similarity 70.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32
Db 23 SQHWSYGWLP 32

RESULT 2
Q9W7G0 PRELIMINARY; PRT; 33 AA.
AC Q9W7G0;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LH-RH) (FRAGMENT).
GN GNRH2.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
RT differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110993; AAD3463.1; -.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3668 MW; 099C825E4A7A2A3BB CRC64;

Query Match 24.3%; Score 41; DB 13; Length 33;
Best Local Similarity 70.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32
Db 23 SQHWSYGWLP 32

RESULT 3
O50832 PRELIMINARY; PRT; 30 AA.
AC O50832;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 3.6 KDA PROTEIN.
GN BBK29.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp36.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31.
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
```

```
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kervatage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
RL EMBL; AE000788; AAC66163.1; -.
DR TIGR; BBK29; -.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 30 AA; 3575 MW; B36F1C321118A2B6 CRC64;

Query Match 21.3%; Score 36; DB 16; Length 30;
Best Local Similarity 35.3%; Pred. No. 3.2e+02;
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 DEKIAKMEKASSVFNV 17
Db 9 DREKMKIRKLSSYYKI 25

RESULT 4
Q9S909 PRELIMINARY; PRT; 27 AA.
AC Q9S909;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ZEAMATIN=22 KDA ANTIFUNGAL PROTEIN.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC Clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RX MEDLINE=92190627; PubMed=1799695;
RA Vigers A.J., Roberts W.K., Selitrennikoff C.P.;
RT "A new family of plant antifungal proteins.";
RL Mol. Plant Microbe Interact. 4:315-323(1991).
DR HSP; P33679; 1DU5.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; thaumatin; 1.
DR PRINTS; PR00347; THAUMATIN.
DR ProDom; PD001321; Thaumatin; 1.
SQ SEQUENCE 27 AA; 2818 MW; CA9AC7B250FD8E50 CRC64;

Query Match 20.7%; Score 35; DB 10; Length 27;
Best Local Similarity 46.7%; Pred. No. 4e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 13 SVFNVVNSGSLHWS 27
Db 1 AVFTVNVOCPTVWA 15

RESULT 5
Q9S908 PRELIMINARY; PRT; 22 AA.
AC Q9S908;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ANTIFUNGAL PROTEIN.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
```

RP SEQUENCE.
RX MEDLINE=92190627; PubMed=1799695;
RA Vigers A.J., Roberts W.K., Sellitrennikoff C.P.;
RT "A new family of plant antifungal proteins";
RL Mol. Plant Microbe Interact. 4:315-323(1991).
DR HSP; P33679; IDU5.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; thaumatin; 1.
SQ SEQUENCE 22 AA; 2294 MW; 2D8E50C9E270526B CRC64;

Query Match 20.1%; Score 34; DB 10; Length 22;
Best Local Similarity 46.7%; Pred. No. 4.5e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 13 SVFNVNNGPSLHWS 27

Db 1 AVFTVVNRCPTVWA 15

RESULT 6

Q9S8K5 PRELIMINARY; PRT; 26 AA.
AC Q9S8K5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GROUP 5 NEUTRAL PATHOGENESIS-RELATED PROTEIN (FRAGMENT).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE
RX MEDLINE=95040741; PubMed=7952963;
RA Kotwa H., Sato F., Yamada Y.;
RT "Characterization of accumulation of tobacco PR-5 proteins by IEF-immunoblot analysis";
RL Plant Cell Physiol. 35:821-827(1994).
DR HSP; P25871; 1AUN.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; thaumatin; 1.
DR ProDom; PD001321; Thaumatin; 1.
SQ SEQUENCE 26 AA; 2753 MW; 6E997F5939ECA591 CRC64;

Query Match 20.1%; Score 34; DB 10; Length 26;
Best Local Similarity 43.8%; Pred. No. 5.4e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 12 SSVFNVNNGPSLHWS 27

Db 1 SGVFEVHNXPYVWA 16

RESULT 7

Q9QVA3 PRELIMINARY; PRT; 25 AA.
AC Q9QVA3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE TP4 SPERMATID BASIC NUCLEAR PROTEIN (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93094251; PubMed=1460032;
RA Unni E., Weistrich M.L.;
RT "Purification and characterization of the rat spermatid basic nuclear protein TP4";
RL J. Biol. Chem. 267:25359-25363(1992).

SQ SEQUENCE 25 AA; 2996 MW; CA50B083B6177376 CRC64;
Query Match 19.5%; Score 33; DB 11; Length 25;
Best Local Similarity 36.4%; Pred. No. 7.2e+02;
Matches 8; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 10 KASSFNVNNGPSLHWSYGLR 31

Db 2 KDSKVRPKVNVSPYVHFMDFR 23

RESULT 8

Q9QHP5 PRELIMINARY; PRT; 27 AA.
AC Q9QHP5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70) (NS1)] (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.;
RT "The genetic heterogeneity of hypervariable region 1 of the viral genome and the sensitivity of hepatitis C virus to interferon alpha therapy";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF166882; AAD52540.1; -.
DR InterPro; IPR002531; HCV_NSI.
DR Pfam; PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2709 MW; 51F0775913268E25 CRC64;

Query Match 18.9%; Score 32; DB 12; Length 27;
Best Local Similarity 33.3%; Pred. No. 1.1e+03;
Matches 6; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 6 AKMEKASSVFNVNGSPS 23

Db 8 AAHSARSASLFTNCPN 25

RESULT 9

Q95LC4 PRELIMINARY; PRT; 33 AA.
AC Q95LC4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BETA-1 ADRENERGIC RECEPTOR (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21540585; PubMed=11683720;
RA Lee J.H., Zhang W., Moran C.;
RT "Comparative porcine gene mapping relative to human chromosomes 9, 10, 20, and 22";
RL Anim. Genet. 32:313-315(2001).
DR EMBL; AF345500; AAK57992.1; -.
KW Receptor.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3694 MW; 7DBEBDC3E268E150 CRC64;

Query Match 18.9%; Score 32; DB 6; Length 33;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 25 HWSYG 29
Db 11 HWDYG 15

RESULT 10

Q73626 PRELIMINARY; PRT; 28 AA.
AC Q73626;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE REVERSE TRANSCRIPTASE (FRAGMENT).
GN POL.

OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-PATIENT P74, PCR;

RX MEDLINE=96242958; PubMed=8624762;

RA Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;

RT "Evolution of zidovudine resistance-associated genotypes in human immunodeficiency virus type 1-infected patients";

RL J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-PATIENT P74, PCR;

RA Leigh Brown A.J.;

RL Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.

DR EMBL; U45192; AAB04343.1; -.

FT NON_TER 1

SQ SEQUENCE 28 AA; 3167 MW; 1593C39FCC660A79 CRC64;

Query Match 18.6%; Score 31.5; DB 15; Length 28;
Best Local Similarity 30.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 5; Mismatches 6; Indels 3; Gaps 1;

QY 6 AKMEKASSVFNVNGPSLUH 25

Db 5 AELEKEGKISKI---GPKTH 21

RESULT 11

Q9R4C3

AC Q9R4C3 PRELIMINARY; PRT; 18 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE PROTOCATCHUATE 3,4-DIOXYGENASE TYPE I BETA SUBUNIT (EC 1.13.11.3)

OS (FRAGMENT).

OC Agrobacterium tumefaciens.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=358;

RN [1]

RP SEQUENCE.

RX MEDLINE=96337865; PubMed=8772173;

RA Hammer A., Stolz A., Knackmuss H.;

RT "Purification and characterization of a novel type of protocatechuate 3,4-dioxygenase with the ability to oxidize 4-sulfocatechol";

RL Arch. Microbiol. 166:92-100(1996).

SQ SEQUENCE 18 AA; 2008 MW; C7EBD971BBE5BBE9 CRC64;

Query Match 18.3%; Score 31; DB 2; Length 18;
Best Local Similarity 36.4%; Pred. No. 9.9e+02;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 20 SGPSLHWSYGL 30
Db 7 TGTFTWDXGI 17

RESULT 12

Q56150

AC Q56150 PRELIMINARY; PRT; 24 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE HYPOTHETICAL 2.4 KDA PROTEIN (FRAGMENT).

OS Streptomyces violaceus (Streptomyces venezuelae).

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1936;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ISP5230;

RX MEDLINE=96070798; PubMed=7592948;

RA Mosher R.H., Camp D.J., Yang K., Brown M.P., Shaw W.V., Vining L.C.;

RT "Inactivation of chloramphenicol by O-phosphorylation. A novel resistance mechanism in Streptomyces venezuelae ISP5230, a chloramphenicol producer";

RL J. Biol. Chem. 270:27000-27006(1995).

DR EMBL; U09991; AAB36571.1; -.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 24 AA; 2443 MW; 49DDCD3DB5FF0745 CRC64;

Query Match 18.3%; Score 31; DB 2; Length 24;

Best Local Similarity 55.6%; Pred. No. 1.4e+03;

Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 21 GPSLHWSYG 29

Db 3 GPDPHWVG 11

RESULT 13

Q9UD88

AC Q9UD88 PRELIMINARY; PRT; 25 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE APO(A) KRINGLE 4-37, APO(A) KRINGLE TYPE 5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94060120; PubMed=7848387;

RA Pfaffinger D., McLean J., Scanu A.M.;

RT "Amplification of human APO(a) kringle 4-37 from blood lymphocyte DNA";

RL Biochim. Biophys. Acta 1225:107-109(1993).

DR HSSP; P00747; ICRN.

SQ SEQUENCE 25 AA; 2928 MW; BEEDD4C62FA480A8 CRC64;

Query Match 18.3%; Score 31; DB 4; Length 25;

Best Local Similarity 57.1%; Pred. No. 1.4e+03;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 22 PSLHWSY 28

Db 6 PSIRWEY 12

RESULT 14

```
Q9T2T5
ID Q9T2T5 PRELIMINARY; PRT; 33 AA.
AC Q9T2T5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE NADH:UBIQUINONE OXIDOREDUCTASE (COMPLEX I) IRON-SULFUR PROTEIN
DE FRACTION 20 KDA POLYPEPTIDE PEPTIDES T-4/T-7.
OS Bos taurus (Bovine).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=92138662; PubMed=1778979;
RA Masui R., Wakabayashi S., Matsubara H., Hatefi Y.;
RT "The amino acid sequence of the 9 kDa polypeptide and partial amino
RT acid sequence of the 20 kDa polypeptide of mitochondrial
RT NADH:ubiquinone oxidoreductase.";
RL J. Biochem. 110:575-582(1991)
SQ SEQUENCE 33 AA; 3764 MW; B3F96D98EE05A0E7 CRC64;

Query Match 18.3%; Score 31; DB 8; Length 33;
Best Local Similarity 33.3%; Pred. NO. 1.9e+03;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 13 SVFNVNVSGLHWSYGL 30
. : : | | | :
Db 14 SMGSCANGGGYHYYSV 31

RESULT 15
Q9IIP8
ID Q9IIP8 PRELIMINARY; PRT; 27 AA.
AC Q9IIP8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Alberto S.-F.;
RT "Influence of the dynamics of Hepatitis C virus quasiespecies in the
RT histological outcome of liver transplantation.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF221397; AAF77959.1; -.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2987 MW; 04A47C9B4D26C7C2 CRC64;

Query Match 18.0%; Score 30.5; DB 12; Length 27;
Best Local Similarity 54.5%; Pred. NO. 1.8e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 22 PSLHWSYGLRP 32
. : : | |
Db 17 PSFHWG--RP 24
```

Search completed: October 10, 2002, 17:16:42
Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:10:37 : Search time 28 Seconds
(without alignments)
130.908 Million cell updates/sec

Title: US-09-848-834A-12

Perfect score: 169

Sequence: 1 DEKKIAKMEKASSVFNVNSGPSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 289672

Minimum DB seq length: 0

Maximum DB seq length: 33

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	134.5	79.6	33	15 AAR62715	LHRH-containing im
2	95	56.2	33	16 AAR83570	IgE CH4 region con
3	89	52.7	19	22 AAM98951	Vaccine related MH
4	89	52.7	21	10 AAP91504	Sequence of modifi
5	89	52.7	21	16 AAR82586	Plasmodium falcipa
6	89	52.7	21	16 AAR78920	Malaria circumspor
7	89	52.7	21	16 AAR75955	P. falciparum CS p
8	89	52.7	21	16 AAR70912	Malaria circumspor
9	89	52.7	21	17 AAU05612	Circumsporozoite h
10	89	52.7	21	18 AAW35440	T-cell stimulatory
11	89	52.7	21	20 AAY23252	Peptide derived fr

12	89	52.7	21	21 AAY80071	Pathogen derived T
13	89	52.7	21	21 AAY54553	T helper cell (Th)
14	89	52.7	21	21 AAY58777	Unidentified pepti
15	89	52.7	21	22 AAB99706	Plasmodium falcipa
16	89	52.7	21	22 AAG62428	Plasmodium falcipa
17	89	52.7	21	22 AAG84517	Plasmodium falcipa
18	89	52.7	21	22 AAG88269	Plasmodium falcipa
19	89	52.7	21	22 AAG89366	Plasmodium falcipa
20	89	52.7	21	22 AAB84447	Sequence of T help
21	89	52.7	21	22 AAB98457	Plasmodium falcipa
22	89	52.7	22	16 AAR82077	Malaria CST3 prote
23	89	52.7	33	22 AAG63663	Peptide comprising
24	89	52.7	33	22 AAG63516	A peptide which ma
25	85	50.3	18	21 AAY49259	CD4+ T cell epitop
26	82	48.5	21	15 AAR65375	Helper T cell epit
27	81	47.9	21	21 AAY70283	Plasmodium falcipa
28	79	46.7	17	16 AAR78919	Malaria circumspor
29	79	46.7	17	16 AAR70911	Malaria circumspor
30	79	46.7	17	21 AAY99032	HLA class II bindi
31	79	46.7	17	22 AAM98950	Vaccine related MH
32	79	46.7	19	21 AAY99033	HLA class II bindi
33	75	44.4	16	22 AAB46168	Plasmodium sp mala
34	75	44.4	16	22 AAB49067	Malaria CS T3 T-ce
35	75	44.4	17	22 AAM98933	Vaccine related MH
36	75	44.4	17	22 AAM98935	Vaccine related MH
37	75	44.4	17	22 AAM98936	Vaccine related MH
38	75	44.4	17	22 AAM98938	Vaccine related MH
39	75	44.4	17	22 AAM98957	Vaccine related MH
40	75	44.4	17	22 AAM98961	Vaccine related MH
41	75	44.4	17	22 AAM98964	Vaccine related MH
42	74.5	44.1	20	22 AAJ04118	P falciparum CS pr
43	74	43.8	17	22 AAM98956	Vaccine related MH
44	74	43.8	17	22 AAM98962	Vaccine related MH
45	74	43.8	17	22 AAM98963	Vaccine related MH

ALIGNMENTS

RESULT 1
AAR62715
ID AAR62715 standard; peptide; 33 AA.
AC AAR62715;
AC AAR62715;
DT 10-SEP-1995 (first entry)
XX
XX LHRH-containing immunogenic peptide.
DE Helper T cell epitope; universal immune stimulator; invasin; hapten;
XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility;
KW Plasmodium falciparum circumsporozoite.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Domain 1..21
FT /note= "plasmodium falciparum circumsporozoite"
FT /note= "helper T cell epitope"
FT Domain 24..33
FT /note= "LHRH hapten"
XX
XX WO9425060-A.
PN
XX
XX
PD 10-NOV-1994.
XX
XX
XX 28-APR-1994; 94WO-US04832.
XX
XX 27-APR-1993; 93US-0057166.
PR 14-APR-1994; 94US-0229275.
XX
XX (LADD/) LADD A E.

PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX
 PS Claim 8; Page 86; 213pp; English.
 XX
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX
 SQ Sequence 33 AA;
 Query Match 79.6%; Score 134.5; DB 15; Length 33;
 Best Local Similarity 90.3%; Pred. No. 2.1e-13;
 Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 2 EKIKAKMEKASSVFNVNSGSLHWSYGLRP 32
 DB 3 EKIKAKMEKASSVFNVNSGGE-HWSYGLRP 32
 RESULT 2
 AAR83570
 ID AAR83570 standard; peptide; 33 AA.
 AC AAR83570;
 XX
 DT 13-JUN-1996 (first entry)
 DE Ige CH4 region contg. peptide immunogen for treating allergies.
 KW Ige; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 KW vaccine; allergy; antibody; constant heavy chain.
 XX
 OS Synthetic.
 XX
 PN W09526365-A1.
 XX
 PD 05-OCT-1995.
 XX
 PF 24-MAR-1995; 95WO-US03741.
 XX
 PR 25-OCT-1994; 94US-0328912.
 PR 28-MAR-1994; 94US-0218461.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang, CY;
 XX
 DR WPI; 1995-351297/45.
 XX

PT Synthetic peptide-based immunogen contg. Ige CH4 peptide and helper
 PT T cell epitope - useful for eliciting antibody prodn. for allergy
 treatment
 XX
 PS Claim 5; Page 72; 87pp; English.
 XX
 CC AAR82592-R82600 and AAR83560-R83581 are peptide immunogens that are
 CC useful in vaccines for treating allergic reactions. In the immunogens,
 CC an Ige CH4 peptide is attached C-terminally to a series of amino acids
 CC including a helper T cell epitope. The immunogen may also opt. contain
 CC a fatty acid or fatty acid derivative, an invasin domain or alpha-NH2.
 CC The immunogen produces high titres of antibodies to the effector site
 CC in human Ige heavy chain (the CH4 domain peptide) which inhibit mast
 CC cell activation and reduce allergen-induced Ige prodn. The immunogens
 CC may be used in either a radially branching multimeric form or a
 CC linearly arranged monomeric form.
 XX
 SQ Sequence 33 AA;
 Query Match 56.2%; Score 95; DB 16; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 EKIKAKMEKASSVFNVNSG 21
 DB 3 EKIKAKMEKASSVFNVNSG 22
 RESULT 3
 AAM98951
 ID AAM98951 standard; Peptide; 19 AA.
 AC AAM98951;
 XX
 DT 07-DEC-2001 (first entry)
 DE Vaccine related MHC ligand peptide SEQ ID NO:54.
 XX
 KW Glutamic acid; glutamine; vaccine; major histocompatibility complex;
 KW MHC; immunomodulator; antiallergic; endocrine; neuroprotectant;
 KW virucidal; bactericidal; antiparasitic; fungicidal; cytostatic;
 KW medicine; pharmaceutical; immune disorder; immune deficiency;
 KW autoimmune; hypersensitivity; allergy; graft rejection; infection;
 KW hormonal disorder; central nervous system disease; cancer; melanoma;
 KW anti-melanoma vaccine; human immunodeficiency virus.
 XX
 OS Plasmodium malariae.
 XX
 PN W0200170772-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 22-MAR-2001; 2001WO-FR00872.
 XX
 PR 23-MAR-2000; 2000FR-0003711.
 XX
 PA (FABR) FABRE MEDICAMENT SA PIERRE.
 XX
 PI Klingner-Hamour C, Corvaia N, Beck A, Goetsch L;
 XX
 DR WPI; 2001-611470/70.
 XX
 PT Stabilized pharmaceutical containing N-terminal glutamic acid or
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
 PT with strong acid
 XX
 PS Claim 9; Page 39; 149pp; French.
 XX
 CC The present invention describes a pharmaceutical compound (I) that
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue
 CC in the form of an addition salt with a strong, physiologically
 CC acceptable acid (II). Also described are: (a) a pharmaceutical
 CC composition containing at least one (I); (b) a vaccine containing at

CC least one (I) where this is a major histocompatibility complex (MHC)
 CC ligand (Ia); (C) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (C) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
 CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful, in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
 CC allergy, graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (I) viral, bacterial, parasitic
 CC or fungal infections; or (II) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I), e.g. melanoma
 CC and human immunodeficiency virus infection. AAM98898 to AAM99592
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.

XX SQ Sequence 19 AA;
 Query Match 52.7%; Score 89; DB 22; Length 19;
 Best Local Similarity 100.0%; Pred. No. 9.7e-07; Mismatches 0; Indels 0; Gaps 0;
 Matches 19; Conservative 0;

QY 2 EKKIARMEKASSVFNVVNS 20
 |||||
 DB 1 EKKIARMEKASSVFNVVNS 19

RESULT 4

AAP91504
 ID AAP91504 standard; peptide; 21 AA.

XX AC AAP91504;

XX DT 13-MAR-1992 (first entry)

XX DE Sequence of modified Plasmodium circumsporozoite T-cell epitope CS.T3.
 XX DE Circumsporozoite peptide; T-cell epitope; immunogenic composition;
 KW vaccine.

XX OS Plasmodium falciparum.

XX PH Key Location/Qualifiers

FT Misc-difference 1..2 /note= "May be H-Asp-Ile, H-Ile, or H-"

FT FT Misc-difference 19..21 /note= "May be Val-Asn-Ser-OH, Val-Asn-OH,

FT Val-OH or -OH "

XX EP343460-A.

XX PD 29-NOV-1989.

XX PF 12-MAY-1989; 89EP-0108618.

XX PR 24-MAY-1988; 89GB-0012214.

XX PA (HOFF) HOFFMANN-LA ROCHE AG.

XX PI Sinigaglia F;

XX DR WPI; 1989-349561/48.

XX Modified Plasmodium CS peptide - used as a universally recognised
 PT T-cell epitope in vaccines to elicit an immune response against
 PT pathogenic agents

XX PS Claim 1; Page 16; 23pp; English.

XX Peptide CS.T3 (residues 1-21) corresponds to residues 378-398 of the
 CC CS protein from P. falciparum but contains 2 Ala residues in place

CC of the native protein's Cys residues at positions 384 and 389. Also
 CC claimed is AAP91504 (or modified forms, see FT) associated with an
 CC antigenic structure representing a B-cell epitope, pref. a multiple
 CC antigenic peptide, esp. multimers of the repeat sequences NAMP
 CC present in P. falciparum CS protein.

XX SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 19; Conservative 0;

QY 2 EKKIARMEKASSVFNVVNS 20
 |||||

DB 3 EKKIARMEKASSVFNVVNS 21

RESULT 5

AAR82586
 ID AAR82586 standard; peptide; 21 AA.

XX AC AAR82586;

XX DT 13-JUN-1996 (first entry)

XX DE Plasmodium falciparum circumsporozoite helper T cell epitope, PF.

XX KW IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 XX KW vaccine; allergy; antibody; constant heavy chain.

XX OS Plasmodium falciparum.

XX PN WO9526365-A1.

XX PD 05-OCT-1995.

XX PF 24-MAR-1995; 95WO-US03741.

XX PR 25-OCT-1994; 94US-0328912.

XX PR 28-MAR-1994; 94US-0218461.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX DR WPI; 1995-351297/45.

XX Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper
 PT T cell epitope - useful for eliciting antibody prodn. for allergy
 PT treatment

XX PS Claim 3; Page 23; 87pp; English.

XX AAR82571-91 are helper T cell epitopes which can be used in the
 CC preparation of a peptide immunogen that is useful in vaccines for
 CC treating allergic reactions. In the immunogen an IgE CH4 peptide
 CC is attached C-terminally to a series of amino acids including a
 CC helper T cell epitope. The immunogen may also opt. contain a fatty
 CC acid or fatty acid derivative, an invasin domain or alpha-NH2. The
 CC immunogen produces high titres of antibodies to the effector site
 CC in human IgE heavy chain (the CH4 domain peptide) which inhibit mast
 CC cell activation and reduce allergen-induced IgE prodn. The immunogens
 CC may be used in either a radially branching multimeric form or a
 CC linearly arranged monomeric form.

XX SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 16; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 19; Conservative 0;

QY 2 EKKIARMEKASSVFNVVNS 20
 |||||

XX
XX
XX

100

Disclosure; Page 13; 59pp; English.

Disclosure; Page 13; 59pp; English.

CC with beta-2 agonists) and adverse hormonal activities (observed in
CC corticosteroid users).
XX
SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. NO. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVVNS 20
| | | | | | | | | | | | | | | | | | | | |
DB 3 EKKIAKMEKASSVFNVVNS 21

RESULT 10
AAW35440
ID AAW35440 standard; peptide; 21 AA.
XX
AC AAW35440;
XX
XX 22-APR-1998 (first entry)
DT
XX
DE T-cell stimulatory peptide from Plasmodium falciparum.
XX
XX T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
KW scaffold; inhibition; metastasis; wound healing; solid phase.
KW
XX
OS Plasmodium falciparum.
XX
XX W09738011-A1.
XX
XX 16-OCT-1997.
PD
XX
XX 03-APR-1997; 97WO-DE00146.
PF
XX
XX 03-APR-1996; 96DK-0000398.
PR
XX
XX (PEPR-) PEPRESEARCH AS.
PA
XX Heegaard PMH, Jakobsen PH;
PI
XX WPT; 1997-512645/47.
DR
XX
XX Non-dendritic peptide carrier linked to a solid phase - useful as a
PT diagnostic agent and as a scaffold for production of chemical
PT derivatives
XX
XX Claim 30; Page 199; 262pp; English.

A non-dendritic peptide carrier (A) has been developed which is coupled through a linker to a solid phase, forming a complex of (A)-solid phase. Where (A) comprises 10-50 amino acids capable of forming a secondary structure in a benign buffer after liberation from the solid phase, and further the (A)-solid phase complex comprises an immunogenic substance and/or an immune mediator coupled on (A). The present sequence represents a specifically claimed T-cell stimulatory peptide from the present invention. An (A)-solid phase complex can be used as a scaffold for the production of chemical derivatives, characterised by covalently attaching molecules at attachment points. Alternatively (A) is used as a scaffold-peptide for the incorporation into an immunostimulating Complex (Iscom) resulting an (A)-Iscom complex which is used for the chemical coupling of antigenic substances in an aqueous solution by conjugation. (A) derivatised with one or more peptides having fibronectin-, laminin- or vitronectin-like binding activities can be used for the promotion of cell-attachment to plastic surfaces, in particular to inhibit tumour growth and metastasis, and for promotion of wound healing. Also a derivatised (A) can be used for the selection of specifically-binding aptamers or as a diagnostic agent. Such diagnostic-(A) molecules could be used to detect molecules derived from or indicative of pregnancy or of a disease, such as an infectious, autoimmune or cancerous disease.

XX Sequence 21 AA;
SQ

Query Match 52.7%; Score 89; DB 18; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVVNS 20
 |||||
 DB 3 EKKIAKMEKASSVFNVVNS 21

RESULT 11

AAAY23252
 ID AAY23252 standard; peptide: 21 AA.

XX AC AAY23252;

DT 31-AUG-1999 (first entry)

DE Peptide derived from Plasmodium falciparum CS protein.

KW Venezuelan equine encephalitis virus; VEE virus; neoplastic disease;
 KW tumour-associated antigen; cytokine; immunity; cancer; tumour;
 KW CS protein.

XX Plasmodium falciparum.

XX WO9930734-Al.

PN 24-JUN-1999.

XX 14-DEC-1998; 98WO-US25725.

XX 18-DEC-1997; 97US-0068080.

XX (SEAR) SEARLE & CO G D.

XX Hippenmeyer PJ;

XX WPI; 1999-395093/33.

XX Using new Venezuelan equine encephalitis virus vectors

PT Claim 5; Page 24; 40pp; English.

XX The specification describes Venezuelan equine encephalitis (VEE) virus
 CC vectors which can be used to express tumour-associated antigens and
 CC cytokines, and thus induce immunity to cancer. The VEE virus vectors
 CC of the invention can be used prevent, treat, and protect against
 CC primary and metastatic neoplastic diseases, especially tumours such
 CC as lung cancer, breast cancer, ovarian cancer, prostate cancer,
 CC pancreatic cancer, gastric cancer, colon cancer, renal cancer,
 CC bladder cancer, melanoma, hepatoma, sarcoma and lymphoma. The
 CC present sequence is derived from the Plasmodium falciparum CS protein,
 CC and can be fused with the target peptide of the invention to provide a
 CC greater stimulation of the immune system.

XX SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 20; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVVNS 20
 |||||
 DB 3 EKKIAKMEKASSVFNVVNS 21

RESULT 12

AAAY80071
 ID AAY80071 standard; Peptide: 21 AA.

XX AC AAY80071;

XX

DT 15-MAY-2000 (first entry)
 DE Pathogen derived Th epitope SEQ ID NO:78.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergic; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX WO9967293-Al.

PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 for immunization against allergy

XX Claim 11; Page 79; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EKKIAKMEKASSVFNVVNS 20
 |||||
 DB 3 EKKIAKMEKASSVFNVVNS 21

RESULT 13

AAAY54553
 ID AAY54553 standard; peptide: 21 AA.

XX AC AAY54553;

XX 25-APR-2000 (first entry)

XX T helper cell (Th) epitope of Plasmodium falciparum circumsporozoite.

XX Human; CD4 protein; antigenic peptide; CDR2-like domain; apoptosis;
 KW syncytia formation; human immune deficiency virus; HIV binding;
 KW CD4-Class II interaction; immunisation; CD4 surface complex;
 KW immune response; transplant rejection; autoimmune disease; psoriasis;

KW rheumatoid arthritis; systemic lupus erythematosus; ciumsporozoite.
 XX
 OS Plasmodium falciparum.
 PN WO9967294-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US14030.
 XX
 PR 20-JUN-1998; 98US-0100409.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 DR WPI; 2000-160579/14.
 XX
 PT New antigenic peptide from the CDR2 domain of CD4, for immunization
 PT against e.g. human immune deficiency virus -
 PS Claim 11; Page 65; 106pp; English.
 XX
 CC The present sequence represents a broadly reactive promiscuous T helper
 CC cell (Th) epitope derived from Plasmodium falciparum ciumsporozoite. It
 CC is conjugated to antigenic peptides derived from the CDR2-like domain of
 CC the human CD4 protein. These antigenic peptides present neutralising
 CC receptor/co-receptor effector sites of the CDR2-like domain. The
 CC peptides evoke effective antibody responses by having optimised
 CC site-specificity. The induced antibodies block human immune deficiency
 CC virus (HIV) binding and syncytia formation. They may also block
 CC CD4-Class II interactions with other cells, deliver signals to T
 CC cells (inhibiting normal CD4+-mediated immunoregulatory functions) or
 CC induce apoptosis of CD4 cells by simultaneous engagement of T cell
 CC receptors. Conjugates and peptides containing the antigenic peptides are
 CC used for active immunisation to generate antibodies against CD4 surface
 CC complexes, especially to prevent binding of HIV to CD4 and thus HIV
 CC infection, but also to treat undesirable immune responses such as
 CC transplant rejection, or autoimmune diseases (rheumatoid arthritis,
 CC systemic lupus erythematosus or psoriasis). These conjugates produce
 CC high-titre antibodies which are broadly neutralising against primary
 CC isolates from all classes of HIV-1 and of HIV-2.
 XX
 SQ Sequence 21 AA;
 Query Match 52.7%; Score 89; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.le-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 EKKIAMEKASSVFNVNS 20
 Db 3 EKKIAMEKASSVFNVNS 21
 ||||||||||||||||
 RESULT 14
 AAY58777
 ID AAY58777 standard; peptide; 21 AA.
 XX
 AC AAY58777;
 XX
 DT 25-APR-2000 (first entry)
 DE Unidentified peptide.
 XX
 KW Helper T cell; Th epitope; feed additive; growth promotion;
 KW somatostatin.
 XX
 OS Unidentified.
 XX
 PN WO9966950-A1.
 XX
 PD 29-DEC-1999.
 XX

PF 21-JUN-1999; 99WO-US13923.
 XX
 PR 20-JUN-1998; 98US-0100415.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 DR WPI; 2000-160560/14.
 XX
 PT New somatostatin helper T-cell epitope conjugate for raising
 PT anti-somatostatin antibodies to enhance growth rate in animal by
 PT reducing growth inhibitory activity of somatostatin -
 PS Disclosure; Page 53; 59pp; English.
 XX
 CC The present sequence is that of an unidentified peptide of the
 CC invention. The invention relates to peptide compositions (see
 CC AAY58739-66) useful as immunogens for growth promotion in farm
 CC animals. The immunogenic peptides contain helper T cell epitopes
 CC which comprise multiple class II MHC motifs and have somatostatin
 CC at either the C- or N-terminus. They may also include an invasin
 CC domain which acts as a general immune stimulator. The helper T
 CC cell epitopes and the invasin domain enhance the immune response
 CC against the somatostatin self-peptide.
 XX
 SQ Sequence 21 AA;
 Query Match 52.7%; Score 89; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.le-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 EKKIAMEKASSVFNVNS 20
 Db 3 EKKIAMEKASSVFNVNS 21
 ||||||||||||||||
 RESULT 15
 AAB99706
 ID AAB99706 standard; peptide; 21 AA.
 XX
 AC AAB99706;
 XX
 DT 06-SEP-2001 (first entry)
 DE Plasmodium falciparum CS protein fragment 378-398 SEQ ID NO:27.
 XX
 KW Human leukocyte antigen A2 binding peptide; HLA class I A2; CTL;
 KW cytotoxic T-cell lymphocyte; tumour associated antigen; CEA; HER2/neu;
 KW MAGE2; MAGE3; p53; vaccine; cancer; cytostatic; immunomodulator;
 KW immunotherapy; immune response.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO200141741-A1.
 XX
 PD 14-JUN-2001.
 XX
 PF 13-DEC-2000; 2000WO-US34318.
 XX
 PR 13-DEC-1999; 99US-0170448.
 PR 05-APR-2000; 2000US-0543608.
 PR 30-MAY-2000; 2000US-0583200.
 XX
 PA (EPIM-) EPIMUNE INC.
 XX
 PI Fikes J, Sette A, Sidney J, Southwood S, Celis E, Keogh E;
 PI Chesnut R;
 XX
 DR WPI; 2001-381489/40.
 XX
 PT Compositions for use in a vaccine for treating, e.g., breast, lung and
 PT colon cancer comprises at least one peptide that comprises an isolated

PT epitope of a tumor-associated antigen -

XX PS Disclosure; Page 31; 86pp; English.

XX
CC The present invention describes a composition (I) comprising at least
CC one peptide that comprises an isolated, prepared epitope consisting of
CC a sequence selected from 25 short amino acid sequences given in AAB99680
CC to AAB99704. Also described are: (1) a composition (II) comprising one
CC or more peptides, and further comprising at least two epitopes selected
CC from the 25 short amino acid sequences (as above), where each of the one
CC or more peptides comprise less than 50 contiguous amino acids that have
CC 100% identity with a native peptide sequence; and (2) a vaccine
CC composition (III) comprising an epitope selected from the 25 short amino
CC acid sequences (as above) and a pharmaceutical excipient. (I) has
CC cytostatic and immunomodulatory activities and can be used in vaccine
CC production and immunotherapy. The peptide epitope compositions (I)-(II)
CC are useful for monitoring an immune response to a tumour associated
CC antigen or when one or more peptides are combined to create a vaccine
CC (III) that stimulates the cellular arm of the immune system. In
CC particular, the vaccine mediates immune responses against tumours in
CC individuals who bear an allele of the human leukocyte antigen (HLA)-A2
CC supertype and improve the standard of care for patients being treated
CC for breast, colon, or lung cancer. The present sequence represents a
CC Plasmodium falciparum CS protein fragment which is given in the
CC exemplification of the present invention.

XX SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIARMEKASSVFNVNS 20

|||||

Db 3 EKKIARMEKASSVFNVNS 21

Search completed: October 10, 2002, 17:15:19

Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:14:38 ; Search time 25 Seconds
(without alignments)
32.242 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKIAKMEKASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 148310

Minimum DB seq length: 0
Maximum DB seq length: 33

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	134.5	79.6	33	1	US-08-446-692-27
2	134.5	79.6	33	2	US-08-488-351A-27
3	89	52.7	21	1	US-08-186-266-6
4	89	52.7	21	1	US-08-446-692-48
5	89	52.7	21	2	US-08-488-351A-48
6	89	52.7	21	3	US-09-100-409A-54
7	89	52.7	21	5	PCT-US95-02121-97
8	89	52.7	21	5	PCT-US95-13841-20
9	79	46.7	17	5	PCT-US95-02121-96
10	75	44.4	16	2	US-08-817-933A-7
11	73.5	43.5	20	1	US-08-465-167A-20
12	73.5	43.5	20	5	PCT-US92-07218-17
13	70	41.4	15	6	5169933-30
14	66	39.1	30	3	US-09-100-414B-71
15	66	39.1	30	4	US-09-303-323-71
16	65	38.5	30	3	US-09-100-414B-73
17	65	38.5	30	4	US-09-303-323-73
18	63.5	37.6	16	1	US-08-465-167A-19
19	63.5	37.6	16	5	PCT-US92-07218-16
20	62	36.7	21	1	US-08-305-871A-12
21	61	36.1	30	3	US-09-100-414B-68
22	61	36.1	30	4	US-09-303-323-68
23	58	34.3	31	1	US-08-446-692-30
24	58	34.3	31	2	US-08-488-351A-30
25	57.5	34.0	29	3	US-09-100-414B-72
26	57.5	34.0	29	4	US-09-303-323-72
27	57	33.7	20	4	US-09-026-276-26

28	57	33.7	20	4	US-09-026-276-30	Sequence 30, Appl
29	57	33.7	26	1	US-08-446-692-29	Sequence 29, Appl
30	57	33.7	26	2	US-08-488-351A-29	Sequence 29, Appl
31	57	33.7	30	1	US-08-446-692-23	Sequence 23, Appl
32	57	33.7	30	2	US-08-488-351A-23	Sequence 23, Appl
33	56	33.1	20	1	US-07-690-983D-40	Sequence 40, Appl
34	56	33.1	20	4	US-09-026-276-29	Sequence 29, Appl
35	56	33.1	20	4	US-09-026-276-31	Sequence 31, Appl
36	56	33.1	24	1	US-07-690-983D-43	Sequence 43, Appl
37	56	33.1	30	1	US-08-446-692-39	Sequence 39, Appl
38	56	33.1	30	2	US-08-488-351A-39	Sequence 39, Appl
39	56	33.1	30	3	US-09-100-414B-66	Sequence 66, Appl
40	56	33.1	30	4	US-09-303-323-66	Sequence 66, Appl
41	55	32.5	16	1	US-08-188-223-7	Sequence 7, Appl
42	55	32.5	16	4	US-08-968-466-7	Sequence 7, Appl
43	55	32.5	16	4	US-08-478-546B-7	Sequence 7, Appl
44	54	32.0	16	1	US-08-453-588-26	Sequence 26, Appl
45	54	32.0	16	3	US-08-521-079-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-446-692-27
; Sequence 27, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Y1
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic univarsal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-27

Query Match 79.6%; Score 134.5; DB 1; Length 33;
Best Local Similarity 90.3%; Pred. No. 1.8e-13;
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 2 EKIAKMEKASSVFNVNSGSLHWSYGLRP 32
Db 3 EKIAKMEKASSVFNVNSGGE-HWSYGLRP 32

RESULT 2

US-08-488-351A-27
; Sequence 27, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488.351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-27

Query Match 79.6%; Score 134.5; DB 2; Length 33;
Best Local Similarity 90.3%; Pred. No. 1.8e-13;
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 EKKIAKMEKASSVFNVNSGSLHWSYGLRP 32
Db 3 EKKIAKMEKASSVFNVNSGGE-HWSYGLRP 32

RESULT 3

US-08-186-266-6
; Sequence 6, Application US/08186266
; Patent No. 5662907
; GENERAL INFORMATION:
; APPLICANT: KUBO, Ralph T.
; APPLICANT: GREY, Howard M.
; APPLICANT: SETTE, Alessandro
; APPLICANT: CELIS, Esteban

; TITLE OF INVENTION: INDUCTION OF ANTI-TUMOR CYTOTOXIC
; TITLE OF INVENTION: T LYMPHOCYTES IN HUMANS USING
; TITLE OF INVENTION: SYNTHETIC PEPTIDE EPITOPES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,266
; FILING DATE: 25-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/159,339
; FILING DATE: 29-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-50-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Plasmodium falciparum CS
; OTHER INFORMATION: protein at positions 378-398."
US-08-186-266-6

Query Match 52.7%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNS 20
Db 3 EKKIAKMEKASSVFNVNS 21

RESULT 4

US-08-446-692-48
; Sequence 48, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-48

Query Match 52.7%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNS 20
|||||
DB 3 EKKIAMEKASSVFNVNS 21

RESULT 5
US-08-488-351A-48
; Sequence 48, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-48

Query Match 52.7%; Score 89; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNS 20
|||||
DB 3 EKKIAMEKASSVFNVNS 21

RESULT 6
US-09-100-409A-54
; Sequence 54, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-54

Query Match 52.7%; Score 89; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNS 20
|||||
Db 3 EKXIAKMEKASSVFNVNS 21

RESULT 7

PCT-US95-02121-97
; Sequence 97, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Malaria circumsporozoite"
; OTHER INFORMATION: 378-398"
PCT-US95-02121-97

Query Match 52.7%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNS 20
|||||
Db 3 EKXIAKMEKASSVFNVNS 21

RESULT 8

PCT-US95-13841-20
; Sequence 20, Application PC/TUS9513841
; GENERAL INFORMATION:

; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-20

Query Match 52.7%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNS 20
|||||
Db 3 EKXIAKMEKASSVFNVNS 21

RESULT 9

PCT-US95-02121-96
; Sequence 96, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992

;; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..17
; OTHER INFORMATION: /note="Malaria circumsporozoite
; OTHER INFORMATION: 382-398"
PCT-US95-02121-96

Query Match 46.7%; Score 79; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KIAMKAKASSVFNVNS 20
Db 1 KIAMKAKASSVFNVNS 17

RESULT 10
US-08-817-933A-7
; Sequence 7, Application US/08817933A
; Patent No. 5945104
; GENERAL INFORMATION:
; APPLICANT: STANWORTH, DENIS R.
; APPLICANT: LEWIN, IAN V.
; TITLE OF INVENTION: PEPTIDES FOR ANTI-ALLERGY TREATMENT
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5945104th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,933A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9422294.0
; FILING DATE: 04-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 179-23
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4005

;; TELEFAX: (703) 816-4100
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-817-933A-7

Query Match 44.4%; Score 75; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIAKMEKASSVFNV 17
Db 1 EKKIAKMEKASSVFNV 16

RESULT 11
US-08-465-167A-20
; Sequence 20, Application US/08465167A
; Patent No. 5750395
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Sette, Alessandro D.
; APPLICANT: Sidney, John C.
; TITLE OF INVENTION: DNA ENCODING MAGE-1 C-TERMINAL
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES (as amended)
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 98111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,167A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,623
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-60-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-167A-20

Query Match 43.5%; Score 73.5; DB 1; Length 20;
Best Local Similarity 94.7%; Pred. No. 0.00012;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 2 EKKIAKMEKASSVFNVNS 20
Db 1 EKKIAKMEKASSVFNVNS 17

Db 3 EKKIAXM-KASSVFNVNS 20

RESULT 12

PCT-US92-07218-17
; Sequence 17, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/07218
; FILING DATE: 19920826
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14137-26-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US92-07218-17

Query Match 43.5%; Score 73.5; DB 5; Length 20;
Best Local Similarity 94.7%; Pred. No. 0.00012;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 EKKIAXM-KASSVFNVNS 20

||||| |||||||

Db 3 EKKIAXM-KASSVFNVNS 20

RESULT 13

5169933-30
; Patent No. 5169933
; APPLICANT: ANDERSON, DAVID C.; MORGAN, CHARLES JR.; FRITZBERG,
; ALAN R.; NICHOLS, EVERETT J.
; TITLE OF INVENTION: CAVALENTLY-LINKED COMPLEXES AND METHODS
; FOR ENHANCED CYTOTOXICITY AND IMAGING
; NUMBER OF SEQUENCES: 45
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/390,241
; FILING DATE: 07-AUG-1989

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 232,337
; FILING DATE: 15-AUG-1988
; SEQ ID NO: 30:
; LENGTH: 15
5169933-30

Query Match 41.4%; Score 70; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KKIAXMEKASSVFNV 17

||||| |||||||

Db 1 KKIAXMEKASSVFNV 15

RESULT 14

US-09-100-414B-71
; Sequence 71, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-71

Query Match 39.1%; Score 66; DB 3; Length 30;
Best Local Similarity 38.7%; Pred. No. 0.0026;
Matches 12; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 2 EKKIAXMEKASSVFNVNSGSLHWSYGLRP 32

||||| : : : : | |||||

Db 1 KKKIITITRIITITIDGGE--HWSYGLRP 29

RESULT 15

US-09-303-323-71
; Sequence 71, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS

NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-71

Query Match 39.1%; Score 66; DB 4; Length 30;
Best Local Similarity 38.7%; Pred. No. 0.0026;
Matches 12; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 2 EKKIAMEKASSVFNVNSGSLHWSYGLRP 32
:| | | : : : : | | | | | |
Db 1 KKKIITRIITITIDGGE--HWSYGLRP 29

Search completed: October 10, 2002, 21:16:43
Job time : 26 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:57:03 ; Search time 12.5 seconds
(without alignments)
261.363 Million cell updates/sec

Title: US-09-848-834A-13

Perfect score: 174

Sequence: 1 XHWSYGLRPGSGPSLKLLSEIKGIVIRLEGVE 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 7455

Minimum DB seq length: 0

Maximum DB seq length: 34

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	33.3	10	1 RHPGG	gonadoliberin - pi
2	58	33.3	10	1 RUSHG	gonadoliberin - sh
3	54	31.0	10	1 RHAQ1	gonadoliberin I -
4	45	25.9	10	2 A21114	gonadoliberin - ch
5	40	23.0	10	1 RHAQ2	gonadoliberin II -
6	40	23.0	10	1 A61126	gonadoliberin - sp
7	40	23.0	10	2 A46030	gonadoliberin I -
8	40	23.0	10	2 B46030	gonadoliberin II -
9	37	21.3	10	2 A49187	gonadotropin-rela
10	36	20.7	32	2 S30756	genome polypeptide
11	33	19.0	20	2 PS0188	superoxide dismuta
12	31	17.8	21	2 A60225	pyruvate dehydroge
13	31	17.8	21	2 C39543	collagen alpha 3(I
14	30.5	17.5	33	2 PH1736	Ig heavy chain v r
15	30.5	17.5	34	2 PH1746	Ig heavy chain v r
16	30.5	17.5	34	2 PH1749	Ig heavy chain v r
17	30.5	17.5	34	2 PH1749	Ig heavy chain v r
18	30	17.2	18	4 I39461	anti-angiogenesis,
19	30	17.2	21	2 D42762	multicatalytic end
20	30	17.2	30	2 S72626	small-cell-variant
21	30	17.2	32	2 JT0017	ferredoxin [2Fe-2S
22	30	17.2	33	2 JT0022	ferredoxin [2Fe-2S
23	29	16.7	17	4 I51887	hypothetical EWSR1
24	29	16.7	18	2 S09723	2S albumin small c
25	29	16.7	20	2 A39543	collagen alpha 1(I
26	29	16.7	22	2 PQ0070	T-cell receptor be
27	29	16.7	26	2 S78761	ribosomal protein
28	29	16.7	30	1 A1BSAF	thermophilic amino
29	29	16.7	30	2 A49955	protein-tyrosine k

ALIGNMENTS

RESULT 1

RHPGG

gonadoliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Note: The synthetic and natural hormones have the same physicochemical and biologic

R:Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 58; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10

RESULT 2

RUSHG

gonadoliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.-

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

3-dehydroquininate d
finger protein ZNF
hypothetical prote
monamine oxidase
hypothetical prote
superoxide dismuta
T-cell receptor J-
hydrogenase (EC 1.
T-cell receptor be
H-2 class I histoc
gene X protein - h
unidentified QM002
hypothetical prote
Ig heavy chain v r
napin small chain
kinesin light chai

30 29 16.7 34 2 S08196
31 29 16.7 34 2 S33993
32 29 16.7 34 2 H82820
33 28.5 16.4 22 2 I59594
34 28.5 16.4 33 2 E84341
35 28 16.1 24 2 S10618
36 28 16.1 24 2 S40139
37 28 16.1 25 2 S29284
38 28 16.1 25 2 D36889
39 28 16.1 29 2 PS0132
40 28 16.1 31 2 S53192
41 28 16.1 31 2 PN0050
42 28 16.1 34 2 E82100
43 27.5 15.8 33 2 PH1742
44 27 15.5 12 2 S70337
45 27 15.5 12 2 S43170

A:Residues: 1-10 <BUR>
A:Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.053; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSYGLRPG 10

RESULT 3
gonadoliberein I - American alligator
RHAQ1
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of Alligator mississippiensis
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 31.0%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.19; 0; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSYGLRPG 10

RESULT 4
gonadoliberein - chum salmon
A21114
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W. Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match 25.9%; Score 45; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 3.6; 2; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSYGLRPG 10

RESULT 5
gonadoliberein II - American alligator
RHAQ2
N:Alternate names: gonadotropin-releasing hormone II

C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: B60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of Alligator mississippiensis
A:Reference number: A60066; MUID:91352338
A:Accession: B60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.0%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 18; 2; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSYGLRPG 10

RESULT 6
gonadoliberein - spotted ratfish
A61126
N:Alternate names: gonadotropin-releasing hormone
C:Species: Hydrolyagus colliei (Spotted ratfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C:Accession: A61126
R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.-Gen. Comp. Endocrinol. 82, 152-161, 1991
A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a holbrooki
A:Reference number: A61126; MUID:91340067
A:Accession: A61126
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Experimental source: brain
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.0%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 18; 2; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:|:|
Db 2 HWSYGLRPG 10

RESULT 7
gonadoliberein I - spiny dogfish
A6030
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A6030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Petermann Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain
A:Reference number: A6030; MUID:92335300
A:Accession: A6030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 23.0%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 18; 2; Indels 0; Gaps 0;

Matches: 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
||||: ||

Db 2 HWSHGWLPG 10

RESULT 8

B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 23.0%; Score 40; DB 2; Length 10;

Best Local Similarity 66.7%; Pred. No. 18;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
||||: ||

Db 2 HWSHGWLPG 10

RESULT 9

A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A:Title: Primary structure and biological activity of a third gonadotropin-releasing hor
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 21.3%; Score 37; DB 2; Length 10;

Best Local Similarity 55.6%; Pred. No. 49;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
||||: ||

Db 2 HWSHGWLPG 10

RESULT 10

S30756
genome polypeptide - foot-and-mouth disease virus Asia (strain Asia 1) (fragment)
C:Species: Aphthovirus Asia (foot-and-mouth disease virus Asia)
C:Date: 03-May-1994 #sequence_revision 14-Sep-1994 #text_change 26-Aug-1999
C:Accession: S30756
R:Sangar, D.V.; Newton, S.E.; Rowlands, D.J.; Clarke, B.E.
Nucleic Acids Res. 15, 3305-3315, 1987
A:Title: All foot and mouth disease virus serotypes initiate protein synthesis at two se
A:Reference number: S30753; MUID:87203363
A:Accession: S30756
A:Molecule type: genomic RNA
A:Residues: 1-32 <SAN>

A:Cross-references: EMBL:M31578; NID:G210495; PIDN:AAA42658.1; PID:G210496
A:Note: the authors translated the codon CAT for residue 5 as Asp
C:Superfamily: foot-and-mouth disease virus genome polypeptide
C:Keywords: alternative initiators; polypeptide
F:1-32/Product: protein lab (fragment) #status predicted <LAB>
F:29-32/Product: protein lab (fragment) #status predicted <LAB>

Query Match 20.7%; Score 36; DB 2; Length 32;

Best Local Similarity 42.1%; Pred. No. 2.4e+02;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 14 PSLKLLSEIKGVIVHRLEG 32
| | | | | : | : |

Db 9 PLLYALREIKALFLSRTQG 27

RESULT 11

PS0188
superoxide dismutase (EC 1.15.1.1) (Cu-Zn), chloroplast - rice (fragment)
C:Species: Oryza sativa (rice)
C:Date: 05-Mar-1993 #sequence_revision 05-Mar-1993 #text_change 05-Mar-1999
C:Accession: PS0188
R:Kawakami, T.; Tsugita, A.
submitted to JIPID, June 1991
A:Reference number: PS0187
A:Accession: PS0188
A:Molecule type: protein
A:Residues: 1-20 <KAW>
C:Experimental source: leaf

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg

C:Keywords: chloroplast; copper; metalloprotein; oxidoreductase; zinc

Query Match 19.0%; Score 33; DB 2; Length 20;

Best Local Similarity 36.8%; Pred. No. 3.9e+02;

Matches 7; Conservative 7; Mismatches 3; Indels 2; Gaps 1;

QY 15 SLKLLSEIKGVIVHRLEGV 33
: | : : | | | : | | |

Db 2 TIKAVAILKG--THQVEGV 18

RESULT 12

A60225
pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) alpha chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 15-Oct-1999
C:Accession: A60225
R:Lawson, R.; Aitken, A.; Yeaman, S.J.
Biochem. Soc. Trans. 11, 298-299, 1983
A:Title: Primary sequence of the N-terminal region of the alpha-subunit of pyruvate d

A:Reference number: A60225

A:Accession: A60225

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-21 <LAW>

C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphate-

C:Keywords: oxidoreductase

Query Match 17.8%; Score 31; DB 2; Length 21;

Best Local Similarity 63.6%; Pred. No. 7.8e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 21 EIKGVIVHRLE 31
| | | : | | |

Db 8 EIKKCDLHRLE 18

RESULT 13

C39543
collagen alpha 3(IX) chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 23-May-1997

Job time : 12.5 secs

C:Accession: C39543
R:Wu, J.J.; Lark, M.W.; Chun, L.E.; Eyre, D.R.
J. Biol. Chem. 266, 5625-5628, 1991
A:Title: Sites of stromelysin cleavage in collagen types II, IX, X, and XI of cartilage.
A:Reference number: A39543; MUID:91170231
A:Accession: C39543
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-21 <WUA>
C:Superfamily: unassigned collagens

Query Match 17.8%; Score 31; DB 2; Length 21;
Best Local Similarity 71.4%; Pred. No. 7.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 RGGSSGP 14
Db 12 RPPAGP 18

RESULT 14

PHI736
Ig heavy chain V region (clone NP-12-2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PHI736
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PHI675; MUID:93301607
A:Accession: PHI736
A:Molecule type: mRNA
A:Residues: 1-33 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 17.5%; Score 30.5; DB 2; Length 33;
Best Local Similarity 43.8%; Pred. No. 1.5e+03;
Matches 7; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 3 WSYGLRPGS-SGPSLK 17
Db 5 WVKGREPGTKSPMQ 20

RESULT 15

PHI746
Ig heavy chain V region (clone NP-12-12) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PHI746
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PHI675; MUID:93301607
A:Accession: PHI746
A:Molecule type: mRNA
A:Residues: 1-34 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 17.5%; Score 30.5; DB 2; Length 34;
Best Local Similarity 43.8%; Pred. No. 1.6e+03;
Matches 7; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 3 WSYGLRPGS-SGPSLK 17
Db 5 WVKGREPGTKSPMQ 20

Search completed: October 10, 2002, 17:03:23

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:52:48 ; Search time 9 seconds
(without alignments)
146.274 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSGPKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 2222

Minimum DB seq length: 0

Maximum DB seq length: 34

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	54	31.0	10	1	GONL_ALLMI	P37041 alligator m
2	47	27.0	10	1	GONL_CLUPA	P81749 clupea pall
3	45	25.9	10	1	GON3_ONCKE	P20367 oncorhynch
4	40	23.0	10	1	GON2_CHICK	P37043 gallus gall
5	40	23.0	10	1	GONL_SQUAC	P27429 squalus aca
6	37	21.3	10	1	GONL_PETMA	P30948 petromyzon
7	32	18.4	10	1	GONL_CHEPR	P80677 chelyosoma
8	31	17.8	24	1	CT31_LITCI	P81851 litorea cit
9	30	17.2	32	1	FER_PORCR	P18821 porphyridiu
10	30	17.2	33	1	FER_PORAE	P18820 porphyridiu
11	29.5	17.0	24	1	COXI_SHEEP	Q9tr30 oviv aries
12	29	16.7	17	1	PC24_BRANA	P81097 brassica na
13	29	16.7	24	1	FRE4_LITIN	P80223 litorea inf
14	29	16.7	30	1	AMPT_BACST	P00728 bacillus st
15	29	16.7	30	1	HETA_RADMA	P58691 radianthus
16	29	16.7	34	1	Z33B_HUMAN	Q06731 homo sapien
17	28.5	16.4	22	1	AOFA_MOUSE	O64133 mus musculu
18	28	16.1	10	1	GON2_CHEPR	P80678 chelyosoma
19	28	16.1	24	1	SODC_RANCA	P23417 rana catesb
20	27.5	15.8	24	1	AMAA_BACTR	P37356 bacillus th
21	27.5	15.8	29	1	COXK_SHEEP	Q9tr28 oviv aries
22	27	15.5	17	1	UP37_UPEMJ	P82044 uperoleia m
23	27	15.5	20	1	ATP4_SPIOL	P80085 spinacia ol
24	27	15.5	32	1	PHSS_DESBN	P13064 desulfovibr
25	27	15.5	33	1	ACT_DICVI	Q24733 dictyocaulu
26	27	15.5	33	1	YC12_PINTH	P41600 pinus thunb
27	26.5	15.2	27	1	TXA3_ANESU	P01535 anemonia su
28	26	14.9	18	1	CPAX_BOVIN	P22779 bos taurus
29	26	14.9	24	1	PCL1_PACGO	P82421 pachycondyl
30	26	14.9	24	1	PCL2_PACGO	P82422 pachycondyl
31	26	14.9	30	1	TL29_SPIOL	P81833 spinacia ol
32	26	14.9	30	1	Y161_TREPA	O83196 treponema p
33	26	14.9	32	1	TRYP_PENMO	P35050 penaeus mon

34 25.5 14.7 26 1 PCW4_PACGO P82426 pachycondyl
35 25.5 14.7 33 1 PEN3_ADECU P35987 canine aden
36 25 14.4 10 1 GRP_RANRI P23260 rana ridibu
37 25 14.4 15 1 UCO6_MAIZE P80612 zea mays (m
38 25 14.4 21 1 RL21_HALCU P05974 halobacteri
39 25 14.4 23 1 COXJ_ONCMY P80333 oncorhynch
40 25 14.4 23 1 GRP_ONCMY Q9ps30 oncorhynch
41 25 14.4 23 1 SODP_PICAB P29427 picea abies
42 25 14.4 25 1 CR14_LITGI P56229 litoria gil
43 25 14.4 25 1 GRP_SCYCA P09472 scyllorhinu
44 25 14.4 27 1 GRP_CANFA P09889 canis famill
45 25 14.4 27 1 GRP_CHICK P01295 gallus gall

ALIGNMENTS

RESULT 1
GONL_ALLMI
ID GONL_ALLMI STANDARD; PRT; 10 AA.
AC P37041: P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Hormone; Amidation; GnRH; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 31.0%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.04;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
Db 2 HWSYGLQPG 10

RESULT 2
GONL_CLUPA
ID GONL_CLUPA STANDARD; PRT; 10 AA.
AC P81749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH)
DE (Luliberin I).
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;

OC Clupeinae; Clupea.
 OX NCBI_TaxID=30724;
 RN [1]
 RP SEQUENCE, AND FUNCTION.
 RC TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
 Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring."
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;
 Query Match 27.0%; Score 47; DB 1; Length 10;
 Best Local Similarity 77.8%; Pred. No. 0.44;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 III:III II
 Db 2 HWSYGLSPG 10
 III:III II
 RESULT 3
 ID GON3_ONCKE STANDARD; PRT; 10 AA.
 AC P20367; P81751;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III) (LH-
 RH III) (Luliberin III).
 GN GNRH3.
 OS Oncorhynchus keta (Chum salmon), and
 OS Clupea pallasi (Pacific herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8018, 30724;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=O.keta;
 RX MEDLINE=83195140; PubMed=6341999;
 RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
 RT "Characterization of a teleost gonadotropin-releasing hormone."
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
 RN [2]
 RP SEQUENCE, AND FUNCTION.
 RC SPECIES=C.pallasi; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
 Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring."
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR PIR: A21114; A21114.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.

FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;
 Query Match 25.9%; Score 45; DB 1; Length 10;
 Best Local Similarity 77.8%; Pred. No. 0.86;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 III:III II
 Db 2 HWSYGLWPG 10
 III:III II
 RESULT 4
 ID GON2_CHICK STANDARD; PRT; 10 AA.
 AC P37043; P20408; P81750;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GNRH-II)
 DE (LH-RH II) (Luliberin II).
 OS Gallus gallus (Chicken).
 OS Alligator mississippiensis (American alligator),
 OS Squalus acanthias (Spiny dogfish),
 OS Hydrolagus colliel (Spotted ratfish) (Pacific ratfish), and
 OS Clupea pallasi (Pacific herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Chicken; TISSUE=Hypothalamus;
 RX MEDLINE=84222059; PubMed=6427779;
 RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
 RA Matsuo H.;
 RT "Identification of the second gonadotropin-releasing hormone in
 RT chicken hypothalamus: evidence that gonadotropin secretion is
 RT probably controlled by two distinct gonadotropin-releasing hormones
 RT in avian species."
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=A.mississippiensis; TISSUE=Brain;
 RX MEDLINE=91352338; PubMed=1882082;
 RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
 RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure of two forms of gonadotropin-releasing hormone
 RT from brains of the American alligator (Alligator mississippiensis)."
 RL Regul. Pept. 33:105-116(1991).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=S.acanthias; TISSUE=Brain;
 RX MEDLINE=92333300; PubMed=1631133;
 RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
 RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
 RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
 RT dogfish brain provides insight into GNRH evolution."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=H.colliel; TISSUE=Brain;
 RX MEDLINE=91340067; PubMed=1678723;
 RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
 Lee T.;
 RT "Primary structure of gonadotropin-releasing hormone from the brain
 RT of a holoccephalan (ratfish: Hydrolagus colliel)."
 RL Gen. Comp. Endocrinol. 82:152-161(1991).
 RN [5]
 RP SEQUENCE, AND FUNCTION.
 RC SPECIES=C.pallasi; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: A61126; A61126.
DR PIR: B60066; RHAQ2.
DR PIR: A61126; A61126.
DR PIR: B46030; B46030.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;
Query Match 23.0%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 4.7;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 2 HWSHGWLPG 10
RESULT 5
GONL_SQUAC STANDARD; PRT; 10 AA.
AC P27429;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
(Luliberin).
OS Squalus acanthias (Spiny dogfish).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
CC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=923335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
Naborniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
dogfish brain provides insight into Gnrh evolution."
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;
Query Match 23.0%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 4.7;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 2 HWSHGWLPG 10
RESULT 6
GON3_PETMA STANDARD; PRT; 10 AA.
ID GON3_PETMA
AC P30948;

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
(Luliberin III).
OS Petromyzon marinus (Sea lamprey).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
CC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
releasing hormone from lamprey brain."
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237AA1F5A3 CRC64;
Query Match 21.3%; Score 37; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 2 HWSHDWKPG 10
RESULT 7
GONL_CHEPR STANDARD; PRT; 10 AA.
ID GONL_CHEPR
AC P80677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I)
(Luliberin I).
OS Chelyosoma productum.
CC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
CC Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RC MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
and the evolutionary implications."
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -!- MASS SPECTROMETRY: MW=1246.56; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
SQ SEQUENCE 10 AA; 1264 MW; 284B3639DB5AB5A3 CRC64;

```
Query Match      18.4%; Score 32; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 71;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
   ||| :||
Db 2 HWSDFKPG 10

RESULT 8
CT31_LITCI
ID CT31_LITCI STANDARD; PRT; 24 AA.
AC P81851; P81852; P81853;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Citropin 3.1.2 [Contains: Citropin 3.1.1; Citropin 3.1.1].
OS Litoria citropa (Australian blue mountains tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=94770;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=99435977; PubMed=10504394;
RA Wegener K.L., Wabnitz P.A., Carver J.A., Bowie J.H., Chia B.C.S.,
RA Wallace J.C., Tyler M.J.;
RT "Host defence peptides from the skin glands of the Australian blue
RT mountains tree-frog Litoria citropa. Solution structure of the
RT antibacterial peptide citropin 1.1.";
RL Eur. J. Biochem. 265:627-637(1999).
CC -!- TISSUE SPECIFICITY: DORSAL AND SUBMENTAL SKIN GLANDS.
FT PEPTIDE 1 24 CITROPIN 3.1.2.
FT PEPTIDE 1 23 CITROPIN 3.1.1.
FT PEPTIDE 1 22 CITROPIN 3.1.1.
FT SEQUENCE 24 AA; 2614 MW; C9001E295BD0E15D CRC64;

Query Match      17.8%; Score 31; DB 1; Length 24;
Best Local Similarity 33.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
   | | | : : |||
Db 10 LKELTGVIEGIQGV 24

RESULT 9
FER_PORC
ID FER_PORC STANDARD; PRT; 32 AA.
AC P18821;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Ferredoxin (Fragment).
OS Porphyridium cruentum.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Porphyridium.
OX NCBI_TaxID=2793;
RN [1]
RP SEQUENCE.
RA Andrew P.W., Rogers L.J., Haslett B.G., Boulter D.;
RT "Comparative properties of ferredoxins from a marine and freshwater
RT species of Porphyridium.";
RL Phytochemistry 20:1293-1298(1981).
CC -!- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC -!- ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC -!- COFACTOR: BINDS 1 2FE-2S CLUSTER.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
DR PIR; J00022; J00022.
DR HSSP; P00246; 4FXC.
DR InterPro; IPR000564; 2Fe2S_ferredoxin.
DR PROSITE; PS00197; 2Fe2S_FERREDOXIN; PARTIAL.
KW Electron transport; Iron-sulfur; Chloroplast.
FT NON_TER 33
FT SEQUENCE 32 AA; 3641 MW; DB440CF9DBFC867 CRC64;

Query Match      17.2%; Score 30; DB 1; Length 33;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVIV 27
   : |||| : | : |
Db 6 VRLLEAGIDV 17

RESULT 11
COXJ_SHEEP
ID COXJ_SHEEP STANDARD; PRT; 24 AA.
AC Q9TR30;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome c oxidase polypeptide VIIa-liver/heart, mitochondrial
DE (EC 1.9.3.1) (Cytochrome c oxidase subunit VIIa-L) (Fragment).
GN COX7A2 OR COX7AL.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver, and Heart;
RX MEDLINE=96092035; PubMed=8529022;
RA Linder D., Freund R., Kadenbach B.;
```

```
DR PROSITE; PS00197; 2Fe2S_FERREDOXIN; PARTIAL.
KW Electron transport; Iron-sulfur; Chloroplast.
FT NON_TER 32
FT SEQUENCE 32 AA; 3502 MW; D5595A4C89EFD69C CRC64;

Query Match      17.2%; Score 30; DB 1; Length 32;
Best Local Similarity 50.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVIV 27
   : |||| : | : |
Db 5 VRLLEAGIDV 16

RESULT 10
FER_PORAE
ID FER_PORAE STANDARD; PRT; 33 AA.
AC P18820;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Ferredoxin (Fragment).
OS Porphyridium aeruginum.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Porphyridium.
OX NCBI_TaxID=2792;
RN [1]
RP SEQUENCE.
RA Andrew P.W., Rogers L.J., Haslett B.G., Boulter D.;
RT "Comparative properties of ferredoxins from a marine and freshwater
RT species of Porphyridium.";
RL Phytochemistry 20:1293-1298(1981).
CC -!- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC -!- ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC -!- COFACTOR: BINDS 1 2FE-2S CLUSTER.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
DR PIR; J00022; J00022.
DR HSSP; P00246; 4FXC.
DR InterPro; IPR000564; 2Fe2S_ferredoxin.
DR PROSITE; PS00197; 2Fe2S_FERREDOXIN; PARTIAL.
KW Electron transport; Iron-sulfur; Chloroplast.
FT NON_TER 33
FT SEQUENCE 33 AA; 3641 MW; DB440CF9DBFC867 CRC64;

Query Match      17.2%; Score 30; DB 1; Length 33;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVIV 27
   : |||| : | : |
Db 6 VRLLEAGIDV 17

RESULT 11
COXJ_SHEEP
ID COXJ_SHEEP STANDARD; PRT; 24 AA.
AC Q9TR30;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome c oxidase polypeptide VIIa-liver/heart, mitochondrial
DE (EC 1.9.3.1) (Cytochrome c oxidase subunit VIIa-L) (Fragment).
GN COX7A2 OR COX7AL.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver, and Heart;
RX MEDLINE=96092035; PubMed=8529022;
RA Linder D., Freund R., Kadenbach B.;
```



```

RN      SEQUENCE, AND MASS SPECTROMETRY.
RP      TISSUE=Skin secretion;
RX      MEDLINE=97368637; PubMed=9225251;
RA      Raftery M.J., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT      "The structures of the frenatin peptides from the skin secretion of
RL      the giant tree frog Litoria infrarenata.";
RJ      J. Pept. Sci. 2:117-124(1996).
RC      -!- FUNCTION: WIDE SPECTRUM ANTIMICROBIAL PEPTIDE.
CC      -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC      GLANDS.
CC      -!- MASS SPECTROMETRY: MW=2493; METHOD=Fab.
KW      Antibiotic; Amphibian skin.
SQ      SEQUENCE 24 AA; 2495 MW; 6986D063947805A1 CRC64;

Query Match 16.7%; Score 29; DB 1; Length 24;
Best Local Similarity 38.9%; Pred. No. 5.3e+02;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      7 LRFSGSGPSLKLLSEIKG 24
       1 : 1 : 1 : 1 : 1 : 1 :
DB      6 LKKGASDFANALYSIKG 23

RESULT 14
AMPT_BACST STANDARD; PRT; 30 AA.
ID      AC P00728;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      15-DEC-1998 (Rel. 37, Last annotation update)
DE      Thermophilic aminopeptidase I alpha chain (EC 3.4.11.-) (Fragment).
OS      Bacillus steartothermophilus.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group;
OC      Bacillus/Staphylococcus group; Geobacillus.
NCBI_TaxID=1422;
RN      [1]
RN      SEQUENCE.
RP      STRAIN=NCIB 8924;
RX      MEDLINE=74087636; PubMed=4521203;
RA      Stoll E., Ericsson L.H., Zuber H.;
RT      "The function of the two subunits of thermophilic aminopeptidase I.";
RL      Proc. Natl. Acad. Sci. U.S.A. 70:3781-3784(1973).
CC      -!- FUNCTION: METALLOENZYME OF HIGH TEMPERATURE STABILITY AND OF
CC      BROAD SPECIFICITY, RELEASING ALL N-TERMINAL AMINO ACIDS.
CC      -!- SUBUNIT: 12 CHAINS OF TWO DIFFERENT BUT HOMOLOGOUS TYPES, ALPHA
CC      AND BETA, WHICH CAN COMBINE IN VARIOUS RATIOS.
CC      -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M42.
DR      PIR; A00908; AIBSAF.
DR      MEROPS; M42.002; -.
KW      Hydrolase; Aminopeptidase; Metalloprotease.
FT      NON_TER 30 30
SQ      SEQUENCE 30 AA; 3274 MW; D712C9C23E618142 CRC64;

Query Match 16.7%; Score 29; DB 1; Length 30;
Best Local Similarity 60.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      16 LKLLSEIKGV 25
       1 : 1 : 1 : 1 : 1 : 1 :
DB      10 LKALTDKGV 19

RESULT 15
HETA_RADMA STANDARD; PRT; 30 AA.
ID      AC P58691;
DT      01-MAR-2002 (Rel. 41, Created)
DT      01-MAR-2002 (Rel. 41, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      Cytolysin RFX-A (Fragment).
OS      Radiantulus macrodactylus (Sea anemone) (Heteractis macrodactylus).
OC      Eukaryota; Metazoa; Cnidaria; Actinoptera; Zoantharia; Actiniaria;

```

OC Stichodactylidae; Heteractis.
 OX NCBI_TaxID=6120;
 RN [1]
 RP SEQUENCE AND CHARACTERIZATION.
 RX PubMed=10645476;
 RA Monastyrnaia M.M., Zykova T.A., Kozlovskaya E.P.;
 RT "Isolation and characteristics of high molecular weight cytolysins
 from the sea anemone Radianthus macrodactylus.";
 RL Bioorg. Khim. 25:733-741(1999).
 CC -!- FUNCTION: Has both cytolytic and hemolytic activity. Pore forming
 CC protein.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE TENEBROSIN FAMILY.
 KW Hemolysis; Toxin; Transmembrane.
 FT NON_TER 30 30
 SQ SEQUENCE 30 AA; 2898 MW; 4B977131E95D8D71 CRC64;
 Query Match 16.7%; Score 29; DB 1; Length 30;
 Best Local Similarity 46.2%; Pred. No. 6.8e+02;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 12 SGPSLKLLSEIKG 24
 Db 10 AGLGLKILIEVLG 22

Search completed: October 10, 2002, 17:01:38
 Job time : 9 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:54:03 ; Search time 19 Seconds
(without alignments)
309.570 Million cell updates/sec

Title: US-09-848-834A-13

Perfect score: 174

Sequence: 1 XHWSYGLRPGSSGSLKLLSEIKGVIHVHLEGEV 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 16130

Minimum DB seq length: 0

Maximum DB seq length: 34

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	45	25.9	33	13 Q9PT34	Q9PT34 oncorhynch
2	45	25.9	33	13 Q9W7G0	Q9W7G0 oncorhynch
3	38	21.8	20	11 Q9EQX8	Q9eqx8 mus musculu
4	38	21.8	32	5 Q9U340	Q9u340 caenorhabdi
5	36	20.7	32	12 Q66857	Q66857 foot-and-mo
6	34	19.5	23	2 Q9AIL1	Q9aill1 magnetospir
7	32.5	18.7	20	11 Q9QVF6	Q9qvif6 rattus sp.
8	32	18.4	24	10 Q9S937	Q9s937 beta vulgar
9	32	18.4	27	7 Q31210	Q31210 mus musculu
10	32	18.4	31	5 Q9TWK5	Q9twk5 mytilus edu
11	32	18.4	33	6 Q9SLC4	Q9slc4 sus scrofa
12	31	17.8	18	2 Q9R584	Q9r584 rhodobacter
13	31	17.8	27	2 Q9R4M2	Q9r4m2 serratia ma
14	31	17.8	29	3 Q96WS1	Q96ws1 schizosacch
15	30.5	17.5	21	2 Q07840	Q07840 rhodobacter
16	30	17.2	18	4 Q13767	Q13767 homo sapien

17	30	17.2	23	2 Q9R5R3	Q9r5r3 photobacter
18	30	17.2	24	4 Q96H37	Q96h37 homo sapien
19	30	17.2	26	2 Q9ZAA4	Q9zaa4 rhodobacter
20	30	17.2	27	12 Q9IJ77	Q9ij77 hepatitis c
21	30	17.2	27	12 Q9IJ76	Q9ij76 hepatitis c
22	30	17.2	27	12 Q9IJ71	Q9ij71 hepatitis c
23	30	17.2	27	12 Q9IIM4	Q9iim4 hepatitis c
24	30	17.2	30	2 Q45966	Q45966 coxiella bu
25	30	17.2	31	4 Q96C35	Q96c35 homo sapien
26	30	17.2	31	12 Q9LJ10	Q9lj10 tt virus. o
27	29	16.7	15	11 Q9QUZ3	Q9quz3 rattus sp.
28	29	16.7	18	3 Q9P897	Q9p897 emericeia
29	29	16.7	18	13 Q90791	Q90791 gallus gall
30	29	16.7	20	11 Q9QUZ4	Q9quz4 rattus sp.
31	29	16.7	24	4 Q9NQV4	Q9nqv4 homo sapien
32	29	16.7	27	12 Q9QI34	Q9qi34 hepatitis c
33	29	16.7	28	10 Q42364	Q42364 triticum ae
34	29	16.7	30	4 Q9UBS6	Q9ubs6 homo sapien
35	29	16.7	30	8 Q9MS7	Q9msp7 nympheaea od
36	29	16.7	31	8 Q9MS71	Q9ms71 lepocincilis
37	29	16.7	32	4 Q9UN22	Q9un22 homo sapien
38	29	16.7	32	13 P82829	P82829 rana luteiv
39	29	16.7	33	2 P82583	P82583 streptococc
40	29	16.7	34	2 Q05171	Q05171 escherichia
41	29	16.7	34	15 Q71808	Q71808 human immun
42	29	16.7	34	15 Q71810	Q71810 human immun
43	29	16.7	34	16 Q9PGH3	Q9pgh3 xylella fas
44	28.5	16.4	26	12 Q9QRT5	Q9qrt5 hepatitis c
45	28.5	16.4	30	4 Q9UM16	Q9um16 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9PT34	PRELIMINARY:	PRT;	33 AA.
AC	Q9PT34;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)			
DE	(LULIBERIN) (FRAGMENT).			
GN	GNRH1.			
OS	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;			
OC	Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.			
OX	NCBI_TaxID=8022;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9312119; PubMed=10385393;			
RA	Von Schalburg K.R., Sherwood N.M.;			
RT	"Regulation and expression of gonadotropin-releasing hormone gene			
RT	differs in brain and gonads in rainbow trout.";			
RL	Endocrinology 140:3012-3024(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	von Schalburg K.R., Sherwood N.M.;			
RL	Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.			
CC	!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY			
CC	SIMILARITY).			
CC	!- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
DR	EMBL; AF110533; AAD43461.1; -;			
DR	InterPro; IPR002047; AKH.			
DR	InterPro; IPR002012; GnrH.			
DR	Pfam; PF00446; GnrH; 1.			
DR	PROSITE; PS00256; AKH; UNKNOWN_1.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	Annotation; Hormone.			
FT	NON_TER 33			
SQ	SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;			

Query Match 25.9%; Score 45; DB 13; Length 33;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
| | | | | | |
DB 25 HWSYGLWPG 33

RESULT 2
Q9W7G0 PRELIMINARY; PRT; 33 AA.

ID Q9W7G0
AC Q9W7G0
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
GN GNRH2.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
RT differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110993; AAD43463.1; -.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; Gnrh; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;

Query Match 25.9%; Score 45; DB 13; Length 33;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
| | | | | | |
DB 25 HWSYGLWPG 33

RESULT 3
Q9EQX8 PRELIMINARY; PRT; 20 AA.

ID Q9EQX8
AC Q9EQX8
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GLUTAMATE DECARBOXYLASE 65 (FRAGMENT).
GN GAD65.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Makinae K., Kobayashi T., Kobayashi T., Sinkawa H., Sakagami H.,
RA Kondo H., Tashiro F., Miyazaki J., Obata K., Tamura S., Tanigawa Y.;

RT "Structure of the Mouse Glutamate Decarboxylase 65 Gene and its
RT Promoter.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032757; BAB20415.1; -.
FT NON_TER 20
SQ SEQUENCE 20 AA; 1989 MW; 28F544B0E4C3768D CRC64;

Query Match 21.8%; Score 38; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 WSYGLRPGSSGP 14
| | | | | | |
DB 9 WSFGSGDGSADP 20

RESULT 4
Q9U340 PRELIMINARY; PRT; 32 AA.

ID Q9U340
AC Q9U340
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE W06G6.9 PROTEIN.
GN W06G6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
DR EMBL; Z83129; CAB63325.1; -.
SQ SEQUENCE 32 AA; 3775 MW; E49D031288C6806A CRC64;

Query Match 21.8%; Score 38; DB 5; Length 32;
Best Local Similarity 41.7%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSG 13
| | | | | | |
DB 18 HWQAAKPGEWG 29

RESULT 5
Q66857 PRELIMINARY; PRT; 32 AA.

ID Q66857
AC Q66857
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FOOT AND MOUTH DISEASE VIRUS (STRAIN ASIA 1) SEROTYPE PROTEIN RNA, 5'
DE END (FRAGMENT).
OS Foot-and-mouth disease virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=12110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87203363; PubMed=3033601;
RA Sangar D.V., Newton S.E., Rowlands D.J., Clarke B.E.;
RT "All foot and mouth disease virus serotypes initiate protein synthesis
RT at two separate AUGs.";
RL Nucleic Acids Res. 15:3305-3315 (1987).
DR EMBL; M31578; AAA42658.1; -.

```

Db      3 HWXYEIQAKFNSXXSGP 20
|| | :: | | |||

RESULT 8
Q9S937 PRELIMINARY; PRT; 24 AA.
ID Q9S937
AC Q9S937;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE H(+)-TRANSLATING (PYROPHOSPHATE-ENERGIZED) INORGANIC PYROPHOSPHATASE
DE BETA-1 POLYPEPTIDE (EC 3.6.1.1) (FRAGMENT).
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=3555;
RN [1]
RP SEQUENCE.
RX MEDLINE=921179265; PubMed=1311852;
RA Sarafian V., Kim Y., Poole R.J., Rea P.A.;
RT "Molecular cloning and sequence of cDNA encoding the pyrophosphate-
RT energized vacuolar membrane proton pump of Arabidopsis thaliana.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1775-1779(1992).
SQ SEQUENCE 24 AA; 2396 MW; CE19F75ADBEFD43B CRC64;

Query Match 18.4%; Score 32; DB 10; Length 24;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 11 SSGPSLKLL 19
:||||:
Db 15 TSGPSLNL 23

RESULT 9
Q31210 PRELIMINARY; PRT; 27 AA.
ID Q31210
AC Q31210;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC CLASS I H2 GENE (HAPLOTYPE D) (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81210176; PubMed=6786753;
RA Steinmetz M., Frelinger J.G., Fisher D., Hunkapiller T., Pereira D.,
RA Weissman S.M., Uehara H., Nathenson S., Hood L.E.;
RT "three cDNA clones encoding mouse transplacental antigens: homology
RT to immunoglobulin genes.";
RL Cell 24:125-134(1981).
DE EMBL: J00406; AAA39699.1; -.
FT NON_TER 1
SQ SEQUENCE 27 AA; 2670 MW; 9BC8CB8B38801DBB CRC64;

Query Match 18.4%; Score 32; DB 7; Length 27;
Best Local Similarity 42.9%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 5 YGLRPGSGPSLKL 18
:||||:
Db 9 YALAPGQSSDMSL 22

RESULT 10
Q9TWK5 PRELIMINARY; PRT; 31 AA.
ID Q9TWK5
AC Q9TWK5;

```

```
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PROXIMAL COLLAGEN, COL-P.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=95230211; PubMed=7714453;
RA Qian X., Waite J.H.;
RT "Exotic collagen gradients in the byssus of the mussel Mytilus edulis.";
RL J. Exp. Biol. 198;633-644(1995).
DR InterPro: IPR000087; Collagen.
SQ SEQUENCE 31 AA; 2648 MW; B1F7708959101A73 CRC64;

Query Match 18.4%; Score 32; DB 5; Length 31;
Best Local Similarity 71.4%; Pred. No. 1.6e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 PGSSGSPS 15
Db 15 PGSTGPT 21

RESULT 11
Q95LC4
ID Q95LC4 PRELIMINARY; PRT; 33 AA.
AC Q95LC4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BETA-1 ADRENERGIC RECEPTOR (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21540585; PubMed=11683720;
RA Lee J.H., Zhang W., Moran C.;
RT "Comparative porcine gene mapping relative to human chromosomes 9, 10, 20, and 22.";
RL Anim. Genet. 32:313-315(2001).
DR EMBL; AF345500; AAK57992.1; -.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3694 MW; 7DBEBDC3E268E150 CRC64;

Query Match 18.4%; Score 32; DB 6; Length 33;
Best Local Similarity 80.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYG 6
Db 11 HWDYG 15.

RESULT 12
Q9R584
ID Q9R584 PRELIMINARY; PRT; 18 AA.
AC Q9R584;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CYTOCHROME B-561 (FRAGMENT).
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group; Rhodobacter.
OX NCBI_TaxID=1061;
```

```
RN [1]
RP SEQUENCE.
RX MEDLINE=93311992; PubMed=8323277;
RA Bartsch R.G., Caffrey M.S., Van Beeumen J.J., Salamon Z., Tollin G., Meyer T.E., Cusanovich M.A.;
RT "Purification and properties of an unusual membrane-derived cytochrome b-561 from the purple phototrophic bacterium Rhodobacter capsulatus, which is structurally related to the bacteriochlorophyll-binding protein, LHII beta.";
RL Arch. Biochem. Biophys. 304:117-122(1993).
SQ SEQUENCE 18 AA; 1860 MW; 916D00479CD86AC6 CRC64;

Query Match 17.8%; Score 31; DB 2; Length 18;
Best Local Similarity 53.3%; Pred. No. 1.2e+03;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 8 RFGSGSPSLKLLSEI 22
Db 4 KAGPSGLSLKAEFI 18

RESULT 13
Q9R4M2
ID Q9R4M2 PRELIMINARY; PRT; 27 AA.
AC Q9R4M2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE METALLO-BETA-LACTAMASE (FRAGMENT).
OS Serratia marcescens.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE.
RX MEDLINE=95303018; PubMed=7783675;
RA Marumo K., Takeda A., Nakamura Y., Nakaya K.;
RT "Purification and characterization of metallo-beta-lactamase from Serratia marcescens.";
RL Microbiol. Immunol. 39:27-33(1995).
DR InterPro: IPR001018; Beta_lactam_B.
DR ProDom: PD007656; Beta_lactam_B; 1.
SQ SEQUENCE 27 AA; 3019 MW; B5FCE3E376DC0A7F CRC64;

Query Match 17.8%; Score 31; DB 2; Length 27;
Best Local Similarity 46.7%; Pred. No. 1.9e+03;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 14 PSLKLLSEIKGVIVH 28
Db 5 PDLKIEKLDEGVVHV 19

RESULT 14
Q96WS1
ID Q96WS1 PRELIMINARY; PRT; 29 AA.
AC Q96WS1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MEU24 (FRAGMENT).
GN MEU24.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD16-1.
RX MEDLINE=21270454; PubMed=11376151;
RA Watanabe T., Miyashita K., Saito T., Yoneki T., Kakiyama Y., Nabeshima K., Kishi Y.A., Shimoda C., Nojima H.;
```

RT "Comprehensive isolation of meiosis-specific genes identifies novel
 RT proteins and unusual non-coding transcripts in Schizosaccharomyces
 RT pombe";
 RL Nucleic Acids Res. 29:2327-2337(2001).
 DR EMBL; AB054308; BAB60875.1; -.
 FT NON_TER 1
 SQ SEQUENCE 29 AA; 3269 MW; F8C8692D0782770B CRC64;

Query Match 17.8%; Score 31; DB 3; Length 29;
 Best Local Similarity 36.0%; Pred. No. 2e+03;
 Matches 9; Conservative 6; Mismatches 8; Indels 2; Gaps 1;

QY 7 LRPSSGPSKLLSEIKGVIVHRLE 31
 I:|:|:| I: : : | :|:|
 DB 2 LKPGSC--SVDWIARILKAVYFLE 24

RESULT 15

O07840
 ID O07840 PRELIMINARY; PRT; 21 AA.
 AC O07840;
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE UV ENDONUCLEASE (FRAGMENT).
 GN UVRA.
 OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 CC Rhodobacter.
 OX NCBI_TaxID=1063;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2.4.1;
 RX MEDLINE=98326818; PubMed=9563685;
 RA Fernandez de Henestrosa A.R., Rivera E., Tapias A., Barbe J.;
 RT "Identification of the Rhodobacter sphaeroides SOS box.";
 RL Mol. Microbiol. 28:991-1003(1998).
 DR EMBL; AF003108; AAC46417.1; -.
 KW Endonuclease.
 FT NON_TER 21
 SQ SEQUENCE 21 AA; 2400 MW; B14ADAC2262DF490 CRC64;

Query Match 17.5%; Score 30.5; DB 2; Length 21;
 Best Local Similarity 33.3%; Pred. No. 1.7e+03;
 Matches 6; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 17 KLLSEIKGVIVHRLEGVE 34
 I:|:|:| I:|:|:|
 DB 5 KFIS-VRGAREHNLRKGD 21

Search completed: October 10, 2002, 17:02:38
 Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:51:33 ; Search time 23 Seconds
(without alignments)
164.196 Million cell updates/sec

Title: US-09-848-834A-13

Perfect score: 174

Sequence: 1 XHWSYGLRPGSGPSLKLSEIKGVIVHRLGVE 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 294572

Minimum DB seq length: 0

Maximum DB seq length: 34

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	44.8	18	AAW35441	T-cell stimulatory
2	77	44.3	25	AAW62705	LHRH-containing im
3	74	42.5	20	AAW57161	Measles virus mimo
4	73.5	42.2	34	AAW05620	mige2-GG-MVFlth.
5	72	41.4	15	AAW62697	Helper T cell epit
6	72	41.4	15	AAW82591	Measles virus F pr
7	72	41.4	15	AAW82082	Measles virus glyc
8	72	41.4	15	AAW05604	Measles virus F pr
9	72	41.4	15	AAW88401	Measles virus F pr
10	72	41.4	15	AAW88392	Measles virus F pr
11	72	41.4	15	AAW80054	Pathogen derived T

12	72	41.4	15	21	AAW91121	Measles virus F pr
13	72	41.4	15	21	AAW44762	Measles virus prot
14	72	41.4	15	21	AAW68540	Helper T cell epit
15	72	41.4	15	21	AAW54537	T helper cell (Th)
16	72	41.4	15	21	AAW58764	Measles virus F pr
17	72	41.4	15	22	AAW84440	Amino acid sequenc
18	72	41.4	15	22	AAW68638	HER-2 B cell pepti
19	72	41.4	27	15	AAW62707	LHRH-containing im
20	72	41.4	27	16	AAW83584	IgE CH4 region con
21	72	41.4	27	16	AAW83576	IgE CH4 region con
22	72	41.4	27	21	AAW91156	MVF Th epitope/LHR
23	72	41.4	27	21	AAW68567	Peptide immunogen
24	72	41.4	34	17	AAW05619	Myf1H-GG-mige25.
25	71	40.8	17	16	AAW78283	GNRH immunomimic a
26	71	40.8	17	21	AAW58140	Gonadotropin relea
27	71	40.8	17	22	AAW93519	Immunomimic peptid
28	69.5	39.9	29	21	AAW91258	Modified MVF Th ep
29	69	39.7	15	21	AAW91128	Modified measles v
30	69	39.7	16	20	AAW67575	T-cell epitope pep
31	69	39.7	19	21	AAW91137	Modified measles v
32	69	39.7	27	21	AAW91163	Modified MVF Th ep
33	69	39.7	31	21	AAW91175	Modified MVF Th ep
34	69	39.7	31	21	AAW91203	Modified MVF Th ep
35	69	39.7	31	21	AAW91206	Modified MVF Th ep
36	68.5	39.4	29	21	AAW91260	Modified MVF Th ep
37	66.5	38.2	29	21	AAW91264	Modified MVF Th ep
38	66	37.9	19	21	AAW91139	Modified measles v
39	66	37.9	31	21	AAW91179	Modified MVF Th ep
40	64	36.8	13	16	AAW82578	Measles virus F pr
41	64	36.8	15	21	AAW91126	Modified measles v
42	64	36.8	15	21	AAW91130	Modified measles v
43	64	36.8	15	21	AAW68544	Helper T cell epit
44	64	36.8	19	21	AAW91135	Modified measles v
45	64	36.8	19	21	AAW68551	Helper T cell epit

ALIGNMENTS

RESULT 1
AAW35441
ID AAW35441 standard; peptide; 18 AA.
XX
AC AAW35441;
XX
DT 22-APR-1998 (first entry)
XX
DE T-cell stimulatory peptide from measles virus F-protein 286-302.
XX
KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
XX scaffold; inhibition; metastasis; wound healing; solid phase.
OS Measles virus.
XX
PN WO9738011-Al.
XX
PD 16-OCT-1997.
XX
PF 03-APR-1997; 97WO-DE00146.
XX
PR 03-APR-1996; 96DK-0000398.
XX
(PEPR-) PEPESEARCH AS.
XX
PI Heegaard PMH, Jakobsen PH;
XX
DR WPI; 1997-512645/47.
XX
PT Non-dendritic peptide carrier linked to a solid phase - useful as a
XX diagnostic agent and as a scaffold for production of chemical
XX derivatives
XX
PS Claim 30; Page 199; 262pp; English.

XX	CC	The present sequence represents a measles virus mimotope for inhibiting
XX	CC	binding of BH129 peptide. The peptide has an amino acid sequence
XX	CC	corresponding to, or mimicking at least 1 antigenic determinant of the
XX	CC	morbillivirus haemagglutinin protein. The peptide can be used to provoke
XX	CC	an immune response to measles virus in mammals, and is useful in the
XX	CC	preparation of a vaccine for the prevention of measles. The peptide is
XX	CC	present in a vaccine at about 0.1 mu g to 100 mu g.
XX	CC	
XX	CC	Sequence 20 AA;
XX	CC	
XX	CC	Query Match 42.5%; Score 74; DB 19; Length 20;
XX	CC	Best Local Similarity 88.9%; Pred. No. 0.00062;
XX	CC	Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	16	LKLLSEIKGVIVHRLEGV 33
DB	3	LSQLSEIKGVIVHRLEGV 20
RESULT 4		
AAW05620	ID	AAW05620 standard; peptide; 34 AA.
XX	AC	AAW05620;
XX	XX	10-DEC-1996 (first entry)
XX	DE	mIge2-GG-WVFlth.
XX	XX	
KW	KW	Immunoglobulin; IgF; membrane protein; human; epsilon chain; hepatitis B;
KW	KW	membrane anchoring domain; helper T cell; surface antigen; core antigen;
KW	KW	pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;
KW	KW	Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;
KW	KW	diphtheria toxin; plasmodium falciparum; circumsporozoite; E. coli TraT;
KW	KW	schistosoma mansoni; triose phosphate isomerase; allergenic reaction;
KW	KW	allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;
KW	KW	antihistamine; decongestant; beta-2 agonist; immunosuppression;
XX	XX	corticosteroid.
OS	OS	Synthetic.
XX	XX	
FH	Key	Location/Qualifiers
FT	Peptide	1..17
FT	FT	/note= "Immunoglobulin E fragment"
FT	FT	20..34
FT	FT	/note= "measles virus F protein helper T cell epitope #1"
XX	XX	
PN	XX	W09612740-A1.
XX	XX	
PD	XX	02-MAY-1996.
XX	XX	
PF	XX	25-OCT-1995; 95WO-US13841.
XX	XX	
PR	XX	25-OCT-1994; 94US-0328519.
XX	XX	
PA	XX	(UNBI-) UNITED BIOMEDICAL INC.
XX	XX	
PI	XX	Walfield AM, Wang CY;
XX	XX	
DR	XX	WPI; 1996-230555/23.
XX	XX	
PT	XX	Peptide immunogen useful in treatment of allergy - comprises
PT	XX	membrane-bound IgE epsilon-chain peptide synthesised linearly in
XX	XX	tandem with T helper epitope peptide
XX	XX	
XX	XX	Claim 3; Page 26; 53pp; English.
XX	XX	
CC	CC	AAW05617-W05624 represent peptide immunogens of the invention. These
CC	CC	sequences are composed of a membrane-bound immunoglobulin E (IgE)
CC	CC	fragment, and a helper T cell epitope. The helper T cell epitopes used
CC	CC	in these peptides can be selected from (amongst others) the hepatitis B
CC	CC	surface or core antigen, pertussis toxin, tetanus toxin, measles virus F

protein, or Chlamydia trachomatis major outer membrane protein. This sequence contains the IGE epsilon heavy chain membrane anchoring domain represented by AA05596, and the measles virus F protein helper T cell epitope shown in AAW05604. The peptide immunogens of the invention can be used in vaccines for the immunotherapeutic treatment of allergic reactions, including allergic rhinitis, food allergies, anaphylaxis, or virally-induced asthma. The immunogens overcome the short effective period of antihistamines, decongestants, and beta-2 agonists, while preventing the broad immunosuppression of corticosteroids. The peptides do not have the potential side effects of restlessness or sedation (associated with antihistamines), associated increased morbidity in asthmatics (as seen with beta-2 agonists) and adverse hormonal activities (observed in corticosteroid users).

Query Match 42.2%; Score 73.5; DB 17; Length 34;
Best Local Similarity 64.3%; Pred. No. 0.0014;
Matches 18; Conservative 1; Mismatches 4; Indels 5; Gaps

Qy 6 GLRPGSSGPSLKLSEIKGVIVHRLEGV 33
I : I | ||||| |||||
Db 12 GEAPWTGG----LSEIKGVIVHRLEGV 34

RESULT 5
AAR62697
ID AAR62697 standard; peptide; 15 AA.
AC AAR62697;
XX XX
DT 10-SEP-1995 (first entry)
XX XX
DE Helper T cell epitope for use in universal immune stimulator.
XX XX
KW Helper T cell epitope; universal immune stimulator; invasin; haptens;
vaccine; measles.
XX XX
OS Measles virus.
XX XX
PN W09425060-A.
XX XX
PD 10-NOV-1994.
XX XX
PF 28-APR-1994; 94WO-US04832.
XX XX
PR 27-APR-1993; 93US-0057166.
PR 14-APR-1994; 94US-0229275.
XX XX
PA (LADD/) LADD A E.
PA (WANG/) WANG C Y.
PA (ZAMB/) ZAMB T.
XX XX
PI Ladd AE, Wang CY, Zamb T;
XX XX
DR WPI; 1994-357910/44.
XX XX
PT Immunogenic luteinising hormone releasing hormone peptide(s) -
PT that suppress LHRR activity in males and females
XX XX
PS Claim 7; Page 25; 213pp; English.
XX XX
CC Synthetic immunogenic peptides are provided in which a universal immune
stimulator is linked to a peptide or protein hapten containing B cell
and/or cytotoxic T lymphocyte epitopes, giving a product which causes
potent immune responses to the coupled peptide or protein. The
stimulator consists of (A) a promiscuous helper T cell epitope (Th)
which elicits an immune response to the coupled peptide in members of
a heterogeneous population expressing diverse HLA phenotypes, and (B)
an adjuvant peptide sequence from the invasin protein of fersinia.
Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
invasin and Th domains and between the immune stimulator and haptens
components. When the haptens is LHRR, then optionally the invasin domain

CC can be omitted from the immune stimulator component.
 CC The present sequence represents a measles virus F protein helper T cell
 CC epitope which can be used as Th in the immune stimulator.

XX SQ Sequence 15 AA;

Query Match 41.4%; Score 72; DB 15; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 |||||
 Db 1 LSEIKGVIVHRLEGV 15

RESULT 6

AAR82591
 ID AAR82591 standard; peptide; 15 AA.

XX AC AAR82591;

XX DT 13-JUN-1996 (first entry)

XX DE Measles virus F protein cell epitope, MVF1.

XX KW IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 KW vaccine; allergy; antibody; constant heavy chain.

XX OS Measles virus.

XX PN WO9526365-A1.

XX PD 05-OCT-1995.

XX PF 24-MAR-1995; 95WO-US03741.

XX PR 25-OCT-1994; 94US-0328912.

XX PR 28-MAR-1994; 94US-0218461.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX DR WPI; 1995-351297/45.

XX Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper
 PT T cell epitope - useful for eliciting antibody prodn. for allergy
 PT treatment

XX PS Claim 3; Page 22; 87pp; English.

XX AAR82571-91 are helper T cell epitopes which can be used in the
 CC preparation of a peptide immunogen that is useful in vaccines for
 CC treating allergic reactions. In the immunogen an IgE CH4 peptide
 CC is attached C-terminally to a series of amino acids including a
 CC helper T cell epitope. The immunogen may also opt. contain a fatty
 CC acid or fatty acid derivative, an invasion domain or alpha-NH2. The
 CC immunogen produces high titres of antibodies to the effector site
 CC in human IgE heavy chain (the CH4 domain peptide) which inhibit mast
 CC cell activation and reduce allergen-induced IgE prodn. The immunogens
 CC may be used in either a radially branching multimeric form or a
 CC linearly arranged monomeric form.

XX SQ Sequence 15 AA;

Query Match 41.4%; Score 72; DB 16; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 |||||
 Db 1 LSEIKGVIVHRLEGV 15

RESULT 7

AAR82082
 ID AAR82082 standard; peptide; 15 AA.

XX AC AAR82082;

XX DT 03-APR-1996 (first entry)

XX DE Measles virus glycoprotein F 288-302 T cell epitope.

XX KW Retro-inverso modified; T cell epitope; analogue; vaccine; B cell;
 KW immunisation; MVF; glycoprotein F; gpF; 288-302.

XX OS Measles virus.

XX PN WO9523166-A1.

XX PD 31-AUG-1995.

XX PF 24-FEB-1995; 95WO-AU00090.

XX PR 25-FEB-1994; 94AU-0004119.

XX PA (DEAK-) DEAKIN RES LTD.

XX PI Comis A, Fischer P, Tyler MI;

XX DR WPI; 1995-311503/40.

XX Synthetic peptide T cell epitope analogue (retro-)inverso modified
 PT - used conjugated with a B cell epitope in a vaccine tailored to a
 PT specific condition, e.g. polio, hepatitis B, etc.

XX PS Claim 3; Page 38; 51pp; English.

XX A vaccine comprising a retro-inverso modified, native T cell epitope
 CC i.e. AAR82073-R82094, conjugated to a B-cell epitope (i.e. the malaria
 CC immunodominant B-cell epitope (AAR82095), and epitopes based on
 CC poliomyelitis, tetanus or hepatitis B proteins etc.) is tailored
 CC to a condition of interest (depending on the epitopes used), and
 CC used to immunise a host against the specific condition.

XX SQ Sequence 15 AA;

Query Match 41.4%; Score 72; DB 16; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 |||||
 Db 1 LSEIKGVIVHRLEGV 15

RESULT 8

AAW05604
 ID AAW05604 standard; peptide; 15 AA.

XX AC AAW05604;

XX DT 10-DEC-1996 (first entry)

XX DE Measles virus F protein helper T cell epitope #1.

XX Immunoglobulin; IgE; membrane protein; human; epsilon chain; hepatitis B;
 KW membrane anchoring domain; helper T cell; surface antigen; core antigen;
 KW pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;
 KW Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;
 KW diphtheria toxin; plasmidium falciparum; circumsporozoite; E. coli Trif;
 KW schistosoma mansoni; triose phosphate isomerase; allergenic reaction;
 KW allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;
 KW antihistamine; decongestant; beta-2 agonist; immunosuppression;
 KW corticosteroid.

XX OS Measles virus.
 XX PN W09612740-A1.
 XX XX
 XX PD 02-MAY-1996.
 XX XX
 XX PF 25-OCT-1995; 95WO-US13841.
 XX XX
 XX PR 25-OCT-1994; 94US-0328519.
 XX XX
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX XX
 XX PI Walfield AM, Wang CY;
 XX XX
 XX DR WPI; 1996-230555/23.
 XX XX
 XX PT Peptide immunogen useful in treatment of allergy - comprises
 PT membrane-bound IgE epsilon-chain peptide synthesised linearly in
 PT tandem with T helper epitope peptide
 XX XX
 XX PS Claim 2; Page 18; 53pp; English.
 XX XX
 XX CC AAW05957-W05616 represent helper T cell epitopes used in the peptide
 CC immunogens of the invention. This sequence represents a measles virus F
 CC protein helper T cell antigen. The peptides of the invention contain
 CC one of these sequences, and a membrane-bound immunoglobulin E (IgE)
 CC fragment (see AAW05595 and AAW05596). The peptide immunogens of the
 CC invention can be used in vaccines for the immunotherapeutic treatment of
 CC allergic reactions, including allergic rhinitis, food allergies,
 CC anaphylaxis, or virally-induced asthma. The immunogens overcome the
 CC short effective period of antihistamines, decongestants, and beta-2
 CC agonists, while preventing the broad immunosuppression of
 CC corticosteroids. The peptides do not have the potential side effects of
 CC restlessness or sedation (associated with antihistamines), associated
 CC adverse morbidity in asthmatics (as seen with beta-2 agonists) and
 CC adverse hormonal activities (observed in corticosteroid users).
 XX XX
 XX SQ Sequence 15 AA;
 Query Match 41.4%; Score 72; DB 17; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 19 LSEIKGVIVHRLEGV 33
 |||||
 Db 1 LSEIKGVIVHRLEGV 15
 RESULT 9
 AAR88401
 ID AAR88401 standard; Peptide; 15 AA.
 XX AC AAR88401;
 XX XX
 XX DT 22-JUN-1996 (first entry)
 XX XX
 XX DE Measles virus F protein MVF peptide B-cell antigen.
 XX XX
 XX KW B-antigen; vaccine; antibody; T-cell; T-lymphocyte;
 XX KW alpha-helix; coiled-coil heterodimer; core peptide; subunit.
 XX OS Synthetic.
 XX XX
 XX PN W09531480-A1.
 XX XX
 XX PD 23-NOV-1995.
 XX XX
 XX PF 18-MAY-1995; 95WO-CA00293.
 XX XX
 XX PR 18-MAY-1994; 94US-0245507.
 XX XX
 XX PA (SPIS-) SPI SYNTHETIC PEPTIDES INC.

XX Cachia PJ, Hodges RS, Houston ME, Irvin RT, Kay CM;
 PI Zhou NE;
 XX WPI; 1996-010880/01.
 XX DR
 XX XX
 XX PT Hetero:dimeric polypeptide immunogen in coiled-coil configuration
 PT with different antigens on each sub:unit - useful in vaccines and
 PT for antibody prodn.
 XX XX
 XX PS Disclosure; Page 63; 95pp; English.
 XX XX
 XX CC This antigenic peptide may be attached to a core peptide
 CC contained in one of the 2 subunits of an alpha-helical coiled-coil
 CC heterodimer. The other subunit may contain e.g. a T-cell antigen
 CC peptide. Each core peptide is comprised of terminal and
 CC internal AA repeat sequences. This peptide antigen is attached
 CC to the core peptide through covalent linkages to certain AA of the
 CC internal repeats. The 2 subunits of the heterodimer are arranged
 CC in a stable alpha-helical coiled-coil configuration having a 1:1
 CC stoichiometry, and the peptide antigen is disposed toward the outer
 CC surfaces of the configuration. The heterodimer may be used as a
 CC synthetic vaccine (optionally multivalent) or to generate
 CC antibodies.
 XX XX
 XX SQ Sequence 15 AA;
 Query Match 41.4%; Score 72; DB 17; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 19 LSEIKGVIVHRLEGV 33
 |||||
 Db 1 LSEIKGVIVHRLEGV 15
 RESULT 10
 AAR88392
 ID AAR88392 standard; Protein; 15 AA.
 XX AC AAR88392;
 XX XX
 XX DT 12-JUN-1996 (first entry)
 XX XX
 XX DE Measles virus F protein MVF-C1.
 XX XX
 XX KW Measles virus; myxo virus; F protein; T-antigen; vaccine; antibody;
 XX KW alpha-helix; coiled-coil heterodimer; core peptide; subunit;
 XX KW plasmid pHIL-S1/E.
 XX OS Synthetic.
 XX XX
 XX PN W09531480-A1.
 XX XX
 XX PD 23-NOV-1995.
 XX XX
 XX PF 18-MAY-1995; 95WO-CA00293.
 XX XX
 XX PR 18-MAY-1994; 94US-0245507.
 XX XX
 XX PA (SPIS-) SPI SYNTHETIC PEPTIDES INC.
 XX XX
 XX PI Cachia PJ, Hodges RS, Houston ME, Irvin RT, Kay CM;
 PI Zhou NE;
 XX WPI; 1996-010880/01.
 XX DR N-PSDB; AAT09872.
 XX XX
 XX PT Hetero:dimeric polypeptide immunogen in coiled-coil configuration
 PT with different antigens on each sub:unit - useful in vaccines and
 PT for antibody prodn.
 XX XX
 XX PS Claim 5; Page 72; 95pp; English.

XX This protein, encoded by DNA cloned in the polylinker region of plasmid
CC PHIL-S1/E, encodes the measles virus F protein MVF-C1 which may be
CC used to form one of the single chain polypeptide subunits of a
CC synthetic vaccine. This antigen may be expressed in tandem with a
CC carrier polypeptide, such that they form a single polypeptide chain,
CC to form a decorated peptide.
XX
SQ Sequence 15 AA;

Query Match 41.4%; Score 72; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 11
AAY80054
ID AAY80054 standard; Peptide; 15 AA.
XX
AC AAY80054;
XX
DT 15-MAY-2000 (first entry)
XX
DE Pathogen derived Th epitope SEQ ID NO:61.
XX
KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS Unidentified.
XX
PN WO9967293-A1.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13959.
XX
PR 20-JUN-1998; 98US-0100287.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY, Walfield AM;
XX
DR WPI; 2000-160578/14.
XX
PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -
XX
PS Claim 11; Page 79; 155pp; English.
XX
CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX

SQ Sequence 15 AA;

Query Match 41.4%; Score 72; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 12
AAY91121
ID AAY91121 standard; peptide; 15 AA.
XX
AC AAY91121;
XX
DT 22-MAY-2000 (first entry)
XX
DE Measles virus F protein promiscuous Th epitope, SEQ ID NO:1.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmidium falci-parum; circumsporozoite; antimalarial; CPTP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Measles virus.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Example 1; Page 54; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope

CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196. AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IGE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 15 AA;

Query Match 41.4%; Score 72; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
 | | | | | | | | | | | | | | | | |
 Db 1 LSEIKGVIVHRLEGV 15

RESULT 13
 AAY44762
 ID AAY44762 standard; peptide; 15 AA.

XX AAY44762;
 AC AAY44762;
 DT 04-MAY-2000 (first entry)
 XX Measles virus protein F derived T-cell activating epitope MVF.
 DE Surface layer protein; S-layer secretion signal; antibiotic; vaccine;
 KW recombinant fusion protein cleavage; enzyme; protein polymer; foodstuff;
 KW antibacterial enzyme; surface glycoprotein; T-cell activating epitope;
 KW MVF; protein F; IPNV; Infectious pancreatic necrosis virus.

XX Measles virus.
 OS
 XX WO200004170-A1.
 PN
 XX 27-JAN-2000.
 PD
 XX 14-JUL-1999; 99WO-CA00637.
 PF
 XX 14-JUL-1998; 98CA-2237704.
 PR
 XX (UYBR-) UNIV BRITISH COLUMBIA.

XX Smit J;
 PI
 XX WPI; 2000-182434/16.

XX Cleavage of Caulobacter produced recombinant fusion proteins useful for
 PT producing vaccine peptides -

XX Example 2; Page 16; 33pp; English.

XX The patent discloses a method for cleaving a recombinant fusion protein
 CC which is produced by Caulobacter and consists of Caulobacter surface
 CC layer (S-layer) protein (containing the C-terminal secretion signal) and
 CC a target protein heterologous to Caulobacter. The cleavage of target
 CC protein from the S-layer protein is carried out under mild acid
 CC conditions so that cleavage occurs at aspartate-proline dipeptide site
 CC without solubilising the protein. The cleavage is accomplished while the
 CC fusion protein is in an insoluble aggregate form which facilitates
 CC purification of the protein. The method is useful for producing pure
 CC proteins including recombinant human and animal therapeutic antibiotic
 CC and vaccine peptides, enzymes, protein polymers, and antibacterial
 CC enzymes for foodstuffs.
 CC The present sequence is a T-cell activating epitope MVF derived
 CC from Measles virus protein F. This sequence was fused to a DNA encoding
 CC a fragment of Infectious pancreatic necrosis virus surface glycoprotein
 CC which is a vaccine candidate. This chimeric protein was in turn fused to
 CC DNA encoding C. crescentus S-layer secretion signal (corresponds to the
 CC C-terminal portion of the S-layer protein from amino acid 690 onwards and
 CC contains native Asp-Pro site) for construction of a recombinant
 CC fusion construct which is expressed in Caulobacter and then cleaved
 CC to recover the vaccine candidate protein.

XX Sequence 15 AA;

Query Match 41.4%; Score 72; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
 | | | | | | | | | | | | | | | | |
 Db 1 LSEIKGVIVHRLEGV 15

RESULT 14
 AAY68540
 ID AAY68540 standard; peptide; 15 AA.

XX AAY68540;
 AC AAY68540;

XX 05-MAY-2000 (first entry)

XX Helper T cell epitope from the F protein of Measles virus.

DE Structured synthetic antigen library; SSAL; helper T cell epitope;
 KW SSAL1 Th1; F protein; Measles virus; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

XX Measles virus.

OS
 XX WO9966952-A1.
 PN
 XX 29-DEC-1999.
 PD
 XX 21-JUN-1999; 99WO-US13960.
 PF
 XX 20-JUN-1998; 98US-0100414.
 PR
 XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;
 PI
 XX WPI; 2000-160562/14.

XX New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer -

PS Example 2; Page 29; 102pp; English.

XX The present sequence represents a helper T cell (Th) epitope of the
 CC F protein of the Measles virus. It was used to design a structured
 CC synthetic antigen library (SSAL), designated SSAL1 Th1. SSAL1 Th1 is
 CC designed to be used in tandem with a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The epitope is used to construct
 CC peptide immunogens of the invention, which contain at least one
 CC antigenic target site, i.e. luteinising hormone-releasing hormone
 CC (LHRH) or its analogue, and an artificial helper T cell epitope (Th).
 CC The peptide immunogens cause induction of a specific immune response
 CC to LHRH which is involved in regulation of spermatogenesis, ovulation,
 CC oestrus, sexual development and secretion of sex hormones. Provision of
 CC a promiscuous T helper epitope (which is functional in genetically
 CC diverse subjects) provides optimum immunogenicity to the B cell
 CC epitopes of the target antigen and thus high antibody titres against
 CC the target antigen. The peptide immunogens of the invention are used
 CC to vaccinate against mammalian LHRH, for use as (reversible)
 CC contraceptive; control of hormone-dependent tumours (cancer of prostate
 CC or breast, also endometriosis); to prevent boar taint (and improve meat
 CC quality) and for immunocastration.

XX Sequence 15 AA;

Query Match 41.4%; Score 72; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 Db | | | | | | | | | | | | | | |
 1 LSEIKGVIVHRLEGV 15

RESULT 15

AA54537
 ID AAY54537 standard; peptide; 15 AA.

XX AC AAY54537;

XX DT 25-APR-2000 (first entry)

XX DE T helper cell (Th) epitope of Measles virus F protein.

XX KW Human; CD4 protein; antigenic peptide; CDR2-like domain; apoptosis;
 KW syncytia formation; human immune deficiency virus; HIV binding;
 KW CD4-Class II interaction; immunisation; CD4 surface complex;
 KW immune response; transplant rejection; autoimmune disease; F protein;
 KW rheumatoid arthritis; systemic lupus erythematosus; psoriasis.

XX OS Measles virus.

XX PN WO9967294-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US14030.

XX PR 20-JUN-1998; 98US-0100409.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX DR WPI; 2000-160579/14.

XX PT New antigenic peptide from the CDR2 domain of CD4, for immunization
 PT against e.g. human immune deficiency virus -

XX PS Claim 11; Page 65; 106pp; English.

XX CC The present sequence represents a broadly reactive promiscuous T helper
 CC cell (Th) epitope derived from the F protein of the Measles virus. It is
 CC conjugated to antigenic peptides derived from the CDR2-like domain of

CC the human CD4 protein. These antigenic peptides present neutralising
 CC receptor/co-receptor effector sites of the CDR2-like domain. The
 CC peptides evoke effective antibody responses by having optimised
 CC site-specificity. The induced antibodies block human immune deficiency
 CC virus (HIV) binding and syncytia formation. They may also block
 CC CD4-Class II interactions with other cells, deliver signals to T
 CC cells (inhibiting normal CD4+-mediated immunoregulatory functions) or
 CC induce apoptosis of CD4 cells by simultaneous engagement of T cell
 CC receptors. Conjugates and peptides containing the antigenic peptides are
 CC used for active immunisation to generate antibodies against CD4 surface
 CC complexes, especially to prevent binding of HIV to CD4 and thus HIV
 CC infection, but also to treat undesirable immune responses such as
 CC transplant rejection, or autoimmune diseases (rheumatoid arthritis,
 CC systemic lupus erythematosus or psoriasis). These conjugates produce
 CC high-titre antibodies which are broadly neutralising against primary
 CC isolates from all classes of HIV-1 and of HIV-2.

XX Sequence 15 AA;

Query Match 41.4%; Score 72; DB 21; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 Db | | | | | | | | | | | | | | |
 1 LSEIKGVIVHRLEGV 15

Search completed: October 10, 2002, 17:00:59
 Job time : 24 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: October 10, 2002, 16:59:56 ; Search time 11 Seconds
(without alignments)
75.497 Million cell updates/sec

Title: US-09-848-834A-13

Perfect score: 174

Sequence: 1 XHWSYGLRPGSSGSLKLLSEIKGVIVHRLGEV 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 149850

Minimum DB seq length: 0
Maximum DB seq length: 34

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	73.5	42.2	34	PCT-US95-13841-28	Sequence 28, Appl
2	72	41.4	15	US-09-100-414B-1	Sequence 1, Appli
3	72	41.4	15	US-09-100-409A-38	Sequence 38, Appl
4	72	41.4	15	US-09-303-323-1	Sequence 1, Appli
5	72	41.4	15	PCT-US95-13841-12	Sequence 12, Appl
6	72	41.4	27	US-08-446-692-19	Sequence 19, Appl
7	72	41.4	27	US-08-488-351A-19	Sequence 19, Appl
8	72	41.4	27	US-09-100-414B-36	Sequence 36, Appl
9	72	41.4	27	US-09-303-323-36	Sequence 36, Appl
10	72	41.4	34	PCT-US95-13841-27	Sequence 27, Appl
11	71	40.8	17	US-08-188-223-6	Sequence 6, Appli
12	71	40.8	17	US-08-968-466-6	Sequence 6, Appli
13	71	40.8	17	US-08-478-546B-6	Sequence 6, Appli
14	69	39.7	15	US-09-100-414B-8	Sequence 8, Appli
15	69	39.7	15	US-09-303-323-8	Sequence 8, Appli
16	69	39.7	16	US-08-460-502-3	Sequence 3, Appli
17	69	39.7	19	US-09-100-414B-17	Sequence 17, Appl
18	69	39.7	19	US-09-303-323-17	Sequence 17, Appl
19	69	39.7	27	US-09-100-414B-43	Sequence 43, Appl
20	69	39.7	27	US-09-303-323-43	Sequence 43, Appl
21	69	39.7	31	US-09-100-414B-55	Sequence 55, Appl
22	69	39.7	31	US-09-100-414B-86	Sequence 86, Appl
23	69	39.7	31	US-09-100-414B-89	Sequence 89, Appl
24	69	39.7	31	US-09-303-323-55	Sequence 55, Appl
25	69	39.7	31	US-09-303-323-86	Sequence 86, Appl
26	69	39.7	31	US-09-303-323-89	Sequence 89, Appl
27	68	39.1	15	US-08-446-692-9	Sequence 9, Appli

28	68	39.1	15	2	US-08-488-351A-9	Sequence 9, Appli
29	68	39.1	25	1	US-08-446-692-17	Sequence 17, Appl
30	68	39.1	25	2	US-08-488-351A-17	Sequence 17, Appl
31	66	37.9	19	3	US-09-100-414B-19	Sequence 19, Appl
32	66	37.9	19	4	US-09-303-323-19	Sequence 19, Appl
33	66	37.9	31	3	US-09-100-414B-59	Sequence 59, Appl
34	66	37.9	31	4	US-09-303-323-59	Sequence 59, Appl
35	64	36.8	15	3	US-09-100-414B-6	Sequence 6, Appli
36	64	36.8	15	3	US-09-100-414B-10	Sequence 10, Appl
37	64	36.8	15	4	US-09-303-323-6	Sequence 6, Appli
38	64	36.8	15	4	US-09-303-323-10	Sequence 10, Appl
39	64	36.8	19	3	US-09-100-414B-15	Sequence 15, Appl
40	64	36.8	19	4	US-09-303-323-15	Sequence 15, Appl
41	64	36.8	27	3	US-09-100-414B-41	Sequence 41, Appl
42	64	36.8	27	3	US-09-100-414B-47	Sequence 47, Appl
43	64	36.8	27	4	US-09-303-323-41	Sequence 41, Appl
44	64	36.8	27	4	US-09-303-323-47	Sequence 47, Appl
45	64	36.8	31	3	US-09-100-414B-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
PCT-US95-13841-28
; Sequence 28, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-28

Query Match 42.2%; Score 73.5; DB 5; Length 34;
Best Local Similarity 64.3%; Pred. No. 0.00029;
Matches 18; Conservative 1; Mismatches 4; Indels 5; Gaps 1;
QY 6 GLRPGSSGSLKLLSEIKGVIVHRLGEV 33

Query Match 41.4%; Score 72; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 5
PCT-US95-13841-12
; Sequence 12, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-12

Query Match 41.4%; Score 72; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 6
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

Query Match 41.4%; Score 72; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 7
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C. H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-19

Query Match 41.4%; Score 72; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LSEIKGVIVHRLEGV 33
DB 1 LSEIKGVIVHRLEGV 15

RESULT 8
US-09-100-414B-36
; Sequence 36, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-36

Query Match 41.4%; Score 72; DB 3; Length 27;

Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LSEIKGVIVHRLEGV 33
DB 1 LSEIKGVIVHRLEGV 15

RESULT 9
US-09-303-323-36
; Sequence 36, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-36

Query Match 41.4%; Score 72; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LSEIKGVIVHRLEGV 33
DB 1 LSEIKGVIVHRLEGV 15

RESULT 10
PCT-US95-13841-27
; Sequence 27, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York

STATE: NY
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13841
FILING DATE: 25-OCT-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/328,519
FILING DATE: 25-OCT-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lin, Maria C.H.
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4117
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-13841-27

Query Match 41.4%; Score 72; DB 5; Length 34;
Best Local Similarity 100.0%; Pred. No. 0.00048;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLGV 33
|||||
Db 1 LSEIKGVIVHRLGV 15

RESULT 11
US-08-188-223-6
; Sequence 6, Application US/08188223
; Patent No. 568506
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scibienski, Robert
; TITLE OF INVENTION: Immunogens Against Gonadotropin
; TITLE OF INVENTION: Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,223
; FILING DATE: 27-JAN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286

TELEFAX: 212-354-8113
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /label= pGlu
; OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoprolinone)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..10
; OTHER INFORMATION: /note= "immunomimic"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 11..17
; OTHER INFORMATION: /note= "spacer"
US-08-188-223-6

Query Match 40.8%; Score 71; DB 1; Length 17;
Best Local Similarity 92.3%; Pred. No. 0.0003;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSGP 14
|||||
Db 2 HWSYGLRPGSSGP 14

RESULT 12
US-08-968-466-6
; Sequence 6, Application US/08968466
; Patent No. 6132720
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scibienski, Robert
; TITLE OF INVENTION: Immunogens Against Gonadotropin
; TITLE OF INVENTION: Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,466
; FILING DATE: 27-JAN-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286
; TELEFAX: 212-354-8113
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FEATURE:

```
/ NAME/KEY: Modified-site
/ LOCATION: 1
/ OTHER INFORMATION: /label= pglu
/ OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoprolinamide)"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 1..10
/ OTHER INFORMATION: /note= "immunomimic"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 11..17
/ OTHER INFORMATION: /note= "spacer"
/
US-08-968-466-6
```

```
Query Match 40.8%; Score 71; DB 4; Length 17;
Best Local Similarity 92.3%; Pred. No. 0.0003;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 HWSYGLRPGSSGP 14
    |||||
Db 2 HWSYGLRPGSSPP 14
```

```
RESULT 13
US-08-478-546B-6
; Sequence 6, Application US/08478546B
; Patent No. 6303123
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scribner, Robert
; TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent
; TITLE OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,546B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/188,223
; FILING DATE: 27-JAN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286
; TELEFAX: 212-354-8113
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified site
; LOCATION: 1
; OTHER INFORMATION: /note= Xaa
; OTHER INFORMATION: /note= "pyroglutamic acid"
; FEATURE:
; NAME/KEY: Region
```

```
/ LOCATION: 1..10
/ OTHER INFORMATION: /note= "immunomimic"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 11..17
/ OTHER INFORMATION: /note= "spacer"
/
US-08-478-546B-6
```

```
Query Match 40.8%; Score 71; DB 4; Length 17;
Best Local Similarity 92.3%; Pred. No. 0.0003;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 HWSYGLRPGSSGP 14
    |||||
Db 2 HWSYGLRPGSSPP 14
```

```
RESULT 14
US-09-100-414B-8
; Sequence 8, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-09-100-414B-8
```

```
Query Match 39.7%; Score 69; DB 3; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.0005;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 19 LSEIKGVIVHRLGV 33
    |||||
Db 1 LSEIKGVIVHKLGV 15
```

```
RESULT 15
US-09-303-323-8
; Sequence 8, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
```

; TITLE OF INVENTION: IMMUNOGENS
 ; NUMBER OF SEQUENCES: 106
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morgan & Finnegan, L.L.P.
 ; STREET: 345 Park Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10154-0054
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC Windows
 ; SOFTWARE: Word 97
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/303,323
 ; FILING DATE: 30-APR-1999
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/100,414
 ; FILING DATE: 20-JUNE-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maria H. Lin
 ; REGISTRATION NUMBER: 29,323
 ; REFERENCE/DOCKET NUMBER: 1151-4157
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-758-4800
 ; TELEFAX: 212-751-6849
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 15 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-303-323-8

Query Match 39.7%; Score 69; DB 4; Length 15;
 Best Local Similarity 93.3%; Pred. No. 0.0005;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 19 LSEIKGVIVHRLQGV 33
 DB 1 LSEIKGVIVHRLQGV 15

Search completed: October 10, 2002, 17:04:06
 Job time : 12 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model
Run on: October 10, 2002, 21:19:02 ; Search time 43 seconds
(without alignments)
82,682 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSGSLFNFTVFWLVPKVSASHLE 37

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 8317

Minimum DB seq length: 0
Maximum DB seq length: 37

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	58	29.0	10	1 RHPGG	gonadoliberin - pi
2	58	29.0	10	1 RHSHG	gonadoliberin - Sh
3	54	27.0	10	1 RHAQ1	gonadoliberin I -
4	45	22.5	10	2 A21114	gonadoliberin - ch
5	40	20.0	10	1 RHAQ2	gonadoliberin II -
6	40	20.0	10	1 A61126	gonadoliberin - sp
7	40	20.0	10	2 A6030	gonadoliberin I -
8	40	20.0	10	2 B46030	gonadoliberin II -
9	40	20.0	30	2 S63531	hypothetical prote
10	37	18.5	10	2 A49187	gonadotropin-relea
11	34.5	17.2	32	2 F71357	hypothetical prote
12	34	17.0	27	2 PN0584	tyrosine 3-monooxy
13	33	16.5	32	2 F87516	hypothetical prote
14	32	16.0	32	2 A56589	galactose binding
15	32	16.0	35	2 S44360	integrin beta-3 -
16	31	15.5	16	2 S57517	r cell receptor be
17	31	15.5	21	2 C39543	collagen alpha 3(I
18	31	15.5	29	2 B43937	endo-1,4-beta-xyla
19	31	15.5	30	2 S09518	prolamin - pearl m
20	30.5	15.2	34	2 PC4167	pigment dispersing
21	30.5	15.2	34	2 C82520	hypothetical prote
22	30	15.0	27	2 PN0583	tyrosine 3-monooxy
23	30	15.0	27	2 PN0585	tyrosine 3-monooxy
24	30	15.0	30	2 S72626	small-cell-variant
25	30	15.0	35	2 C25159	hypothetical prote
26	30	15.0	37	2 D82636	hypothetical prote
27	29.5	14.8	37	2 PH1736	ig heavy chain v r
28	29.5	14.8	34	2 PH1746	ig heavy chain v r
29	29.5	14.8	34	2 PH1747	ig heavy chain v r

30	29.5	14.8	34	2 PH1749	ig heavy chain v r
31	29.5	14.8	36	2 PH1751	ig heavy chain v r
32	29.5	14.8	36	2 PH1752	ig heavy chain v r
33	29.5	14.8	36	2 PH1753	ig heavy chain v r
34	29	14.5	18	2 S09723	2S albumin small c
35	29	14.5	20	2 A39543	collagen alpha 1(I
36	29	14.5	20	2 S27351	lysophospholipase
37	29	14.5	26	2 T11816	hypothetical prote
38	29	14.5	26	2 S78761	ribosomal protein
39	29	14.5	27	2 PN0586	tyrosine 3-monooxy
40	29	14.5	30	2 A49555	protein-tyrosine k
41	28.5	14.2	14	2 S09721	2S albumin small c
42	28.5	14.2	21	2 S48632	carbonic anhydrase
43	28.5	14.2	22	2 PQ0070	T-cell receptor be
44	28.5	14.2	26	2 PQ0779	NADH dehydrogenase
45	28.5	14.2	31	2 A55430	conotoxin NgVIA -

ALIGNMENTS

RESULT 1

RHPGG
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation: synthesis
A:Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 29.0%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. NO. 0.041;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10

RESULT 2

RHSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780
A:Molecule type: protein

A:Residues: 1-10 <BUR>
A:Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormones
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 29.0%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.041; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0

QY 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10
|||||

RESULT 3
RHAQ1
gonadoliberein I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson Regl. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 27.0%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.15; Mismatches 1; Indels 0; Gaps 0;
Matches 8; Conservative 1

QY 2 HWSYGLRPG 10
Db 2 HWSYGLQPG 10
|||||

RESULT 4
A21114
gonadoliberein - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W. Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match 22.5%; Score 45; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 2.8; Mismatches 0; Indels 2; Gaps 0;
Matches 7; Conservative 0

QY 2 HWSYGLRPG 10
Db 2 HWSYGLWPG 10
|||||

RESULT 5
RHAQ2
gonadoliberein II - American alligator
N:Alternate names: gonadotropin-releasing hormone II

C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: B60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson Regl. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: B60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 20.0%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 14; Mismatches 6; Indels 0; Gaps 0;
Matches 6; Conservative 1

QY 2 HWSYGLRPG 10
Db 2 HWSHGWPG 10
|||||

RESULT 6
A61126
gonadoliberein - spotted ratfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Hydroloagus colliel (spotted ratfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C:Accession: A61126
R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T. Gen. Comp. Endocrinol. 82, 152-161, 1991
A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a holostean fish
A:Reference number: A61126; MUID:91340067
A:Accession: A61126
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Experimental source: brain
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 20.0%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 14; Mismatches 6; Indels 0; Gaps 0;
Matches 6; Conservative 1

QY 2 HWSYGLRPG 10
Db 2 HWSHGWPG 10
|||||

RESULT 7
A46030
gonadoliberein I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Petermann, R. Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 20.0%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 14; Mismatches 2; Indels 0; Gaps 0;
Matches 6; Conservative 1

QY 2 HWSYGLRPG 10
Db 2 HWSHGWPG 10
|||||

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||

Db 2 HWSHGWLPG 10

RESULT 8

B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 20.0%; Score 40; DB 2; Length 10;

Best Local Similarity 66.7%; Pred. No. 14;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||

Db 2 HWSHGWYPG 10

RESULT 9

S63531
hypothetical protein 1 - Sulfolobus solfataricus (fragment)
C:Species: Sulfolobus solfataricus
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S63531
R:Jones, C.E.; Fleming, T.M.; Cowan, D.A.; Littlechild, J.A.; Piper, P.W.
Eur. J. Biochem. 233, 800-808, 1995
A:Title: The phosphoglycerate kinase and glyceraldehyde-3-phosphate dehydrogenase genes
expression in Escherichia coli.
A:Reference number: S63528; MUID:96085144
A:Accession: S63531
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-30 <JON>
A:Cross-references: EMBL:X80178

Query Match 20.0%; Score 40; DB 2; Length 30;

Best Local Similarity 62.5%; Pred. No. 50;

Matches 10; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 12 SGPSLFNFTVSWLR 27
|||||

Db 2 SEPSYF----VSEWLR 13

RESULT 10

A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A:Title: Primary structure and biological activity of a third gonadotropin-releasing hor
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A:Status: preliminary
A:Molecule type: protein

A:Residues: 1-10 <SOW>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 18.5%; Score 37; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 38;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||

Db 2 HWSHDWKPQ 10

RESULT 11

F71357
hypothetical protein TP0169 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: F71357
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770
A:Accession: F71357
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-32 <COL>
A:Cross-references: GB:AE001201; GB:AE000520; NID:g3322431; PIDN:AAC65161.1; PID:g332

Query Match 17.2%; Score 34.5; DB 2; Length 32;

Best Local Similarity 38.5%; Pred. No. 3.2e+02;

Matches 10; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 3 WSYGLRPGSGSPSLFNFTVSWLRV 28
|||||

Db 2 WS-GLFPDLQGTAFRAVVASARFRV 26

RESULT 12

PN0584
tyrosine 3-monooxygenase (EC 1.14.16.2) type 3 and 4 - gorilla (fragment)
N:Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase
C:Species: Gorilla gorilla (gorilla)
C:Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000
C:Accession: PN0584
R:Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.
Biochem. Biophys. Res. Commun. 195, 158-165, 1993
A:Title: Increased heterogeneity of tyrosine hydroxylase in humans.
A:Reference number: PN0575; MUID:93371398
A:Accession: PN0584
A:Molecule type: genomic RNA
A:Residues: 1-27 <ICH>
A:Cross-references: GB:L14796
A:Experimental source: lymphocytes of peripheral blood
C:Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine bi

Query Match 17.0%; Score 34; DB 2; Length 27;

Best Local Similarity 38.5%; Pred. No. 3.1e+02;

Matches 10; Conservative 3; Mismatches 9; Indels 4; Gaps 1;

QY 10 GSGSPSLFNFTVSWLRPKVSASH 35
|:|||||

Db 1 GAPGPSL-----TGSLLWPGTAAPAASY 22

RESULT 13

Search completed: October 10, 2002, 22:34:08
Job time : 44 secs

Qy 8 RPSGSPSLFNNFTV 22
| | | | | | | |
| | | | | | | |
Db 4 RIARSPFLFGRFAT 18

Query Match	16.0%;	Score 32;	DB 2;	Length 32;
Best Local Similarity	31.6%;	Pred. No. 7.3e+02;		
Matches	6;	Conservative	5;	Mismatches 8;
				Indels 0;
				Gaps 0;

Query Match 16.0%; Score 32; DB 2; Length 35;
Best Local Similarity 47.4%; Pred. No. 8e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 4; Gaps 2;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:17:23 ; Search time 25 seconds
(without alignments)
57.305 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHWSYGLRPGSGSPSFLNFTVSEWLRVPKVSASHLE 37

Scoring Table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 2569

Minimum DB seq length: 0

Maximum DB seq length: 37

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	54	27.0	10	1	GON1_ALLMI	P37041 alligator m
2	47	23.5	10	1	GON1_CLUPA	P81749 clupea pall
3	45	22.5	10	1	GON3_ONCKE	P20367 oncorhynch
4	40	20.0	10	1	GON2_CHICK	P37043 gallus gall
5	40	20.0	10	1	GONL_SQUAC	P27429 squalus aca
6	37	18.5	10	1	GON3_PETWA	P30948 petromyzon
7	34.5	17.2	32	1	Y169_TREPA	O83199 treponema p
8	32	16.0	10	1	GON1_CHEPR	P80677 chelyosoma
9	32	16.0	12	1	UR2_SCYCA	P35490 scyllorhinu
10	30	15.0	31	1	SARL_HUMAN	O00631 homo sapien
11	29.5	14.8	22	1	LANN_STRMU	P80666 streptococc
12	29	14.5	20	1	LP73_HUMAN	P56643 homo sapien
13	28.5	14.2	31	1	CXD6_CONNI	P56710 conus nigro
14	28.5	14.2	36	1	NUC6_SOLTU	P80264 solanum tub
15	28	14.0	10	1	GON2_CHEPR	P80678 chelyosoma
16	28	14.0	31	1	ER29_BOVIN	P81623 bos taurus
17	27	13.5	32	1	ER29_CHICK	P81628 gallus gall
18	27	13.5	32	1	ER29_TRIVU	P81629 trichosurus
19	27	13.5	33	1	PAP1_PARPV	P81865 pardachirus
20	27	13.5	33	1	PAP4_PARMA	P81861 pardachirus
21	27	13.5	33	1	PAP5_PARMA	P81862 pardachirus
22	27	13.5	33	1	PETM_CVAPA	P48366 cyanophora
23	27	13.5	34	1	RNL1_PIG	P15466 sus scrofa
24	26.5	13.2	27	1	TXA3_ANESU	P01535 anemonia su
25	26	13.0	18	1	AG1_EUPMA	P33889 euphorbia m
26	26	13.0	25	1	ACP_ERYLO	P80919 erythroba
27	26	13.0	37	1	PAP3_PARPV	P81866 pardachirus
28	26	13.0	37	1	VGJ_BPPHX	P03651 bacterioph
29	25.5	12.8	30	1	PLMS_SQUAC	P82542 squalus aca
30	25.5	12.8	35	1	FAS_CAPHI	P08757 capra hircu
31	25.5	12.8	37	1	PRF1_RAT	P18889 rattus norv
32	25	12.5	10	1	GRP_RANRI	P23260 rana ridibu
33	25	12.5	15	1	UC06_MAIZE	P80612 zea mays (m

34	25	12.5	21	1	RLZ1_HALCU	P05974 halobacteri
35	25	12.5	22	1	ATP6_COTJA	P50681 coturnix co
36	25	12.5	23	1	GRP_ONCMY	O9ps30 oncorhynch
37	25	12.5	25	1	COXO_ONCMY	P80334 oncorhynch
38	25	12.5	25	1	GRP_SCYCA	P09472 scyllorhinu
39	25	12.5	27	1	GRP_CANFA	P08989 canis fami
40	25	12.5	27	1	GRP_CHICK	P01295 gallus gall
41	25	12.5	27	1	GRP_PIG	P01294 sus scrofa
42	25	12.5	28	1	GRP_ALLMI	P31886 alligator m
43	25	12.5	29	1	PSAF_SYNPG	P31083 synechococc
44	25	12.5	33	1	PAP2_PARPV	P23067 pardachirus
45	25	12.5	33	1	PSBE_SYNVU	P12238 synechococc

ALIGNMENTS

RESULT 1
GON1_ALLMI
ID GON1_ALLMI STANDARD; PRT; 10 AA.
AC P37041: P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 27.0%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.055; 0; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0;

QY 2 HWSYGLRPG 10
|||||:
DB 2 HWSYGLQPG 10

RESULT 2
GON1_CLUPA
ID GON1_CLUPA STANDARD; PRT; 10 AA.
AC P81749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH)
DE (Luliberin I).
GN GnRH1.
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;

OC Clupeinae: Clupea.
OX NCBI_TaxID=30724;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;

Query Match 23.5%; Score 47; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 0.54;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
|||:||||
DB 2 HWSYGLSPG 10

RESULT 3
ID GON3_ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III) (LH-
RH III) (Luliberin III).
GN GNRH3.
OS Oncorhynchus keta (Chum salmon), and
OS Clupea pallasi (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O. keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone."
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C. pallasi; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: A21114; A21114.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

Query Match 22.5%; Score 45; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 1;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLWPG 10

RESULT 4
ID GON2_CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GNRH-II)
(LH-RH II) (Luliberin II).
OS Gallus gallus (Chicken).
OS Alligator mississippiensis (American alligator),
OS Squalus acanthias (Spiny dogfish),
OS Hydrolagus collii (Spotted ratfish) (Pacific ratfish), and
OS Clupea pallasi (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=Chicken; TISSUE=Hypothalamus;
RX MEDLINE=84222059; PubMed=6427779;
RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
RA Matsuo H.;
RT "Identification of the second gonadotropin-releasing hormone in
RT chicken hypothalamus: evidence that gonadotropin secretion is
RT probably controlled by two distinct gonadotropin-releasing hormones
RT in avian species."
RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=A. mississippiensis; TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
RT from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
RN [3]
RP SEQUENCE.
RC SPECIES=S. acanthias; TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
RT dogfish brain provides insight into GNRH evolution."
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN [4]
RP SEQUENCE.
RC SPECIES=H. collii; TISSUE=Brain;
RX MEDLINE=91340067; PubMed=1678723;
RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
RA Lee T.;
RT "Primary structure of gonadotropin-releasing hormone from the brain
RT of a holoccephalan (ratfish: Hydrolagus collii).";
RL Gen. Comp. Endocrinol. 82:152-161(1991).
RN [5]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C. pallasi; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
hormones, including a novel form, from an ancient teleost, herring."
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: B60066; RHAQ2.
DR PIR: A61126; A61126.
DR PIR: B46030; B46030.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;

Query Match 20.0%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 5.4;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:|
DB 2 HWSHGWPG 10

RESULT 5
GONL_SQUAC STANDARD; PRT; 10 AA.
AC P27429;
DT 01-AUG-1992 (Rel. 23; Created)
DT 01-AUG-1992 (Rel. 23; Last sequence update)
DT 15-DEC-1998 (Rel. 37; Last annotation update)
DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
(Luliberin).
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Namvongchon S., Craig A.G.,
RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
dogfish brain provides insight into GNRH evolution."
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;

Query Match 20.0%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 5.4;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:|
DB 2 HWSHGWPG 10

RESULT 6
GON3_PETMA STANDARD; PRT; 10 AA.
ID GON3_PETMA
AC P30948;

DT 01-JUL-1993 (Rel. 26; Created)
DT 01-JUL-1993 (Rel. 26; Last sequence update)
DT 15-DEC-1998 (Rel. 37; Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
(Luliberin III).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
releasing hormone from lamprey brain."
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237A1F5A3 CRC64;

Query Match 18.5%; Score 37; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 14;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||:|
DB 2 HWSHDWPG 10

RESULT 7
Y169_TREPA STANDARD; PRT; 32 AA.
ID Y169_TREPA
AC O83199;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical protein TP0169.
GN TP0169.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
spirochete."
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----

```
DR EMBL; AE001201; AAC65161.1.; -.
KW TIGR; TP0169; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 32 AA; 3616 MW; C01A97771405398F CRC64;

Query Match 17.2%; Score 34.5; DB 1; Length 32;
Best Local Similarity 38.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 3 WSYGLRPGSSGSLNNFTVSFWLRV 28
||| : : : : : ||
2 WS-GLFPDLOGTAFFRAWVASARFV 26

RESULT 8
GONL_CHEPR STANDARD; PRT; 10 AA.
AC P80677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I)
DE (Luliberin I).
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
RT and the evolutionary implications.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996)
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
CC THE GONADUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -!- MASS SPECTROMETRY: MW=1246.56; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro; IPR002012; GNRH.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
FT MOD_RES 10 10 284B3639DBSAB5A3 CRC64;
SQ SEQUENCE 10 AA; 1264 MW; 284B3639DBSAB5A3 CRC64;

Query Match 16.0%; Score 32; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
||| : :
2 HWSDYFKPG 10

RESULT 9
UR2_SCYCA STANDARD; PRT; 12 AA.
AC P35490;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Urotensin II (U-II) (UII).
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
OX NCBI_TaxID=7830;
RN [1]

RP SEQUENCE.
RX TISSUE=Spinal cord;
RC MEDLINE=92319231; PubMed=1620290;
RA Conlon J.M., O'Harte F., Smith D.D., Balmert R.J., Hazon N.;
RT "Purification and characterization of urotensin II and parvalbumin
RT from an elasmobranch fish, Scyllorhinus canicula (common dogfish).";
RL Neuroendocrinology 55:230-235(1992).
CC -!- FUNCTION: HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC COTICOTROPIN-RELEASING FACTOR. PROBABLY INVOLVED IN SMOOTH
CC MUSCLE STIMULATION.
CC -!- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.
DR InterPro; IPR001483; Urotensin_II.
DR Pfam; PF02083; Urotensin_II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11
SQ SEQUENCE 12 AA; 1526 MW; 804729F9D579CEBA CRC64;

Query Match 16.0%; Score 32; DB 1; Length 12;
Best Local Similarity 62.5%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 NNFTVSFW 25
||| : :
1 NNFSDCFW 8

RESULT 10
SARL_HUMAN STANDARD; PRT; 31 AA.
ID SARL_HUMAN
AC 000631;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sarcopolipin.
DE SLN.
GN SLN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98035878; PubMed=9367679;
RA Odermatt A., Taschner P.E.M., Scherer S.W., Beatty B., Khanna V.K.,
RA Cornblath D.R., Chaudhry V., Yee W.-C., Schrank B., Karpati G.,
RA Breuning M.H., Knoers N., MacLennan D.H.;
RT "Characterization of the gene encoding human sarcopolipin (SLN), a
RT proteolipid associated with SERCA1: absence of structural mutations
RT in five patients with Brody disease.";
RL Genomics 45:541-553(1997).
CC -!- SUBUNIT: ASSOCIATED WITH CALCIUM ATPASE SERCA1.
CC -!- SUBCELLULAR LOCATION: MEMBRANE ASSOCIATED. SARCOPLASMIC RETICULUM.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; U96094; AAB86981.1; -.
DR EMBL; U96093; AAB86980.1; -.
DR MIM; 602203; -.
KW Transmembrane; Sarcoplasmic reticulum.
FT TRANSMEM 8 25
SQ SEQUENCE 31 AA; 3762 MW; 9B245D9ACD26C38F CRC64;

Query Match 15.0%; Score 30; DB 1; Length 31;
Best Local Similarity 36.0%; Pred. No. 5.1e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 6; Gaps 1;

QY 10 GSSGSLPNFT-----VSFWLRV 28
```



```

RL Biochem. J. 288:831-837(1992).
CC -!- FUNCTION: DEGRADATION OF LYSOPHOSPHOLIPIDS.
CC -!- CATALYTIC ACTIVITY: 2-lysophosphatidylcholine + H(2)O =
CC glycerophosphocholine + a fatty acid anion.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
KW Hydrolase; Lipid degradation.
FT NON_TER 1
FT NON_TER 20
SQ SEQUENCE 20 AA; 2263 MW; 4D298EGD3F21F87F CRC64;

Query Match 14.5%; Score 29; DB 1; Length 20;
Best Local Similarity 45.5%; Pred. No. 4.3e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 15 SLFNNFTVSF 25
|||:|:|:|
Db 5 SLYNSNTLSVF 15

RESULT 13
CXD6_CONNI CXD6_CONNI STANDARD; PRT; 31 AA.
AC P56710;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Delta-conotoxin NgVIA.
OS Conus nigropunctatus (Black-spot cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=39967;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Venom;
RX MEDLINE=95138099; PubMed=7836370;
RA Fainzilber M., Lodder J.C., Kitz K.S., Kofman O., Vinnitsky I.,
RA Van Rietschoten J., Zlotkin E., Gordon D.;
RT "A new conotoxin affecting sodium current inactivation interacts with
RT the delta-conotoxin receptor site.";
RJ J. Biol. Chem. 270:1123-1129(1995).
CC -!- FUNCTION: INHIBITS SODIUM CHANNEL INACTIVATION. POTENT PARALYTIC
CC TOXIN IN VERTEBRATE AND INVERTEBRATE SYSTEMS. ITS BINDING APPEARS
CC TO BE VOLTAGE-DEPENDENT.
CC -!- SUBCELLULAR LOCATION: Secreted.
KW Venom; Neurotoxin; Sodium channel inhibitor; Hydroxylation.
FT MOD_RES 6 6 HYDROXYLATION.
FT MOD_RES 14 14 BY SIMILARITY.
FT DISULFID 3 18 BY SIMILARITY.
FT DISULFID 10 22 BY SIMILARITY.
FT DISULFID 17 27 BY SIMILARITY.
SQ SEQUENCE 31 AA; 3368 MW; 8387385C693118D1 CRC64;

Query Match 14.2%; Score 28.5; DB 1; Length 31;
Best Local Similarity 35.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 4; Mismatches 8; Indels 1; Gaps 1

Qy 6 GLRPG-SSGSPSLFNNFTVSF 24
|:|:| |:|:|
Db 11 GIKPLGCCVRCFSLFCISF 30

RESULT 14
NUCM_SOLTU NUCM_SOLTU STANDARD; PRT; 36 AA.
AC P80264;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-ubiquinone oxidoreductase 49 kDa subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (NADH dehydrogenase subunit 7) (Complex I-42.5KD) (CI-
DE 42.5KD) (Fragment).
GN NAD7.
OS Solanum tuberosum (Potato).

```

OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC STRAIN=CV, BINTJE; TISSUE=Tuber;
RX MEDLINE=94124587; PubMed=8294484;
RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,
RA Gromann L.;
RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
RT the respiratory chain from the inner mitochondrial membrane of
RT Solanum tuberosum.";
RL J. Biol. Chem. 269:2263-2269(1994).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE. COMPONENT OF THE IRON-SULFUR (IP) FRAGMENT OF
CC THE ENZYME.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- COFACTOR: BINDS 1 4Fe-4S CLUSTER.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC THIS IS A COMPONENT OF THE IRON-SULFUR PROTEIN FRACTION.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 49 kDa SUBUNIT FAMILY.
DR PIR; A49732; A49732.
DR InterPro; IPR001135; Complex1_49Kd.
DR PROSITE; PS00535; COMPLEX1_49K; PARTIAL.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Iron-sulfur; 4Fe-4S.
FT NON_TER 36 36
SQ SEQUENCE 36 AA; 4012 MW; C8A65AA061ACB49F CRC64;

Query Match 14.2%; Score 28.5; DB 1; Length 36;
Best Local Similarity 41.2%; Pred. No. 9.9e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

QY 19 NFTVSWLVRVPKVSASH 35
| | | : | | : | |
Db 9 NFTLNFG--POHPAAH 22

RESULT 15
GON2_CHEPR STANDARD; PRT; 10 AA.
AC P80678;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GNRH-II)
DE (luliberin II).
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
RT and the evolutionary implications.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
CC THE GONADUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -!- MASS SPECTROMETRY: MW=1117.52; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro; IPR002012; Gnrh.
DR PROSITE; PS00473; GNRH; 1.

KW Hormone; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 6 6 INTERCHAIN.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
SQ SEQUENCE 10 AA; 1135 MW; 284B38D1EEB735A3 CRC64;

Query Match 14.0%; Score 28; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
| | | | |
Db 2 HWSLCHAPG 10

Search completed: October 10, 2002, 21:24:39
Job time : 25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 21:16:47 ; Search time 29 Seconds
(without alignments)
220.718 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHWSYGLRPGSGSLFNNFTVSFWRVVKVYASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 19511

Minimum DB seq length: 0

Maximum DB seq length: 37

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertibrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	45	22.5	33	13 Q9PT34	Q9pt34 oncorhynchu
2	45	22.5	33	13 Q9W7G0	Q9w7g0 oncorhynchu
3	40	20.0	31	1 Q55314	Q55314 sulfobolus
4	38	19.0	20	11 Q9EQX8	Q9eqx8 mus musculus
5	38	19.0	32	5 Q9U340	Q9u340 caenorhabdi
6	35	17.5	25	5 Q9BM38	Q9bm38 caenorhabdi
7	35	17.5	29	10 Q9S8N0	Q9s8n0 raphanus sa
8	35	17.5	32	2 Q05602	Q05602 pseudomonas
9	34	17.0	32	6 Q9N1V7	Q9n1v7 equus cabal
10	34	17.0	33	10 Q9XGG9	Q9xgg9 pisum sativ
11	33	16.5	25	4 Q9UQB1	Q9uqb1 homo sapien
12	33	16.5	33	6 Q9SLC4	Q9slc4 sus scrofa
13	33	16.5	33	16 Q9A6D5	Q9a6d5 caulobacter
14	32.5	16.2	20	11 Q9QVF6	Q9qvf6 rattus sp.
15	32	16.0	13	2 Q9L8K1	Q9l8k1 enterococcu
16	32	16.0	22	11 Q9CVJ5	Q9cvj5 mus musculu

17	32	16.0	26	5 Q9BM15	Q9bm15 drosophila
18	32	16.0	31	5 Q9TWK5	Q9twk5 mytilus edu
19	32	16.0	32	5 Q9TWU2	Q9twu2 spodoptera
20	31.5	15.8	25	6 Q9EGJ1	Q9egj1 tarsius ban
21	31.5	15.8	30	4 Q9UM16	Q9umi6 homo sapien
22	31	15.5	13	2 Q9WW72	Q9ww72 enterococcu
23	31	15.5	13	2 Q9WW71	Q9ww71 enterococcu
24	31	15.5	13	2 Q9X5I4	Q9x5i4 enterococcu
25	31	15.5	18	2 Q9LAP4	Q9lap4 enterococcu
26	31	15.5	25	5 Q9BM39	Q9bm39 caenorhabdi
27	31	15.5	29	2 Q9R5M6	Q9r5m6 streptomyce
28	31	15.5	31	11 Q9CQD6	Q9cqdd mus musculu
29	31	15.5	35	15 Q9OQX5	Q9oqx5 human immun
30	30.5	15.2	27	5 Q9BM48	Q9bm48 monilliformi
31	30.5	15.2	34	16 Q9P9M9	Q9p9m9 xyella fas
32	30	15.0	23	2 Q9R5R3	Q9r5r3 photobacter
33	30	15.0	24	4 Q9GH37	Q9gh37 homo sapien
34	30	15.0	26	5 Q9BM51	Q9bm51 monilliformi
35	30	15.0	26	5 Q9BM42	Q9bm42 euperipatoi
36	30	15.0	27	5 Q9BM49	Q9bm49 monilliformi
37	30	15.0	30	2 Q45966	Q45966 coxiella bu
38	30	15.0	33	12 Q9LJ17	Q9lj17 tt virus. o
39	30	15.0	33	12 Q9LJ16	Q9lj16 tt virus. o
40	30	15.0	33	12 Q9LJ15	Q9lj15 tt virus. o
41	30	15.0	33	12 Q9LJ14	Q9lj14 tt virus. o
42	30	15.0	33	12 Q9LJ13	Q9lj13 tt virus. o
43	30	15.0	33	12 Q9LJ12	Q9lj12 tt virus. o
44	30	15.0	33	12 Q9LJ10	Q9lj10 tt virus. o
45	30	15.0	33	12 Q9LJ09	Q9lj09 tt virus. o

ALIGNMENTS

RESULT 1

Q9PT34 PRELIMINARY; PRT; 33 AA.
ID Q9PT34;
AC Q9PT34;
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
GN GNRH1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110533; AAD3461.1; -.
DR InterPro; IPR002047; AKH.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH. 1.
DR PROSITE; PS00256; AKH; UNKNOWN_1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;

```
Query Match      22.5%; Score 45; DB 13; Length 33;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
    ||||| ||
Db 25 HWSYGLWPG 33

RESULT 2
ID Q9W7G0 PRELIMINARY; PRT; 33 AA.
AC Q9W7G0;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
GN GNRH2.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99312119; PubMed=10385393;
RA Von Schalburg K.R., Sherwood N.M.;
RT "Regulation and expression of gonadotropin-releasing hormone gene
RT differs in brain and gonads in rainbow trout.";
RL Endocrinology 140:3012-3024 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA von Schalburg K.R., Sherwood N.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF110993; AAD43463.1; -.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;

Query Match      22.5%; Score 45; DB 13; Length 33;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
    ||||| ||
Db 25 HWSYGLWPG 33

RESULT 3
ID Q55314 PRELIMINARY; PRT; 31 AA.
AC Q55314;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE URF2 PROTEIN (FRAGMENT).
GN URF2.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96085144; PubMed=8521845;
RA Jones C.E., Fleming T.M., Cowan D.A., Littlechild J.A., Piper P.W.;
RT "The phosphoglycerate kinase and glyceraldehyde-3-phosphate
RT dehydrogenase genes from the thermophilic archaeon Sulfolobus
```

```
RT solfataricus overlap by 8bp. Isolation, sequencing of the genes and
RT expression on Escherichia coli.";
RL Eur. J. Biochem. 233:800-808(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94082761; PubMed=8259927;
RA Arcari P., Russo A.D., Iannicello G., Gallo M., Bocchini V.;
RT "Nucleotide sequence and molecular evolution of the gene coding for
RT glyceraldehyde-3-phosphate dehydrogenase in the thermoacidophilic
RT archaeobacterium Sulfolobus solfataricus.";
RL Biochem. Genet. 31:241-251(1993).
DR EMBL; X80178; CAA56461.1; -.
FT NON_TER 31
SQ SEQUENCE 31 AA; 3554 MW; 9A2538F911C7309A CRC64;

Query Match      20.0%; Score 40; DB 1; Length 31;
Best Local Similarity 62.5%; Pred. No. 84;
Matches 10; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 12 SGPSLFNFTVSWLR 27
    ||| |||||
Db 2 SEPSYF----VSWLR 13

RESULT 4
ID Q9EQX8 PRELIMINARY; PRT; 20 AA.
AC Q9EQX8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GLUTAMATE DECARBOXYLASE 65 (FRAGMENT).
GN GAD65.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Makinae K., Kobayashi T., Kobayashi T., Sinkawa H., Sakagami H.,
RA Kondo H., Tashiro F., Miyazaki J., Obata K., Tamura S., Yanagawa Y.;
RT "Structure of the Mouse Glutamate Decarboxylase 65 Gene and Its
RT Promoter.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032757; BAB20415.1; -.
FT NON_TER 20
SQ SEQUENCE 20 AA; 1989 MW; 28F544B0E4C3768D CRC64;

Query Match      19.0%; Score 38; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 WSYGLRPGSSGP 14
    ||:| ||:|
Db 9 WSYGSEDGSADP 20

RESULT 5
ID Q9U340 PRELIMINARY; PRT; 32 AA.
AC Q9U340;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE W06G6.9 PROTEIN.
GN W06G6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
```

RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z83129; CAB63325.1; -
SQ SEQUENCE 32 AA; 3775 MW; E49D031288C6806A CRC64;

Query Match 19.0%; Score 38; DB 5; Length 32;
Best Local Similarity 41.7%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSG 13
||| :|||
Db 18 HWQAAKPGWG 29

RESULT 6
Q9BM38 PRELIMINARY; PRT; 25 AA.
ID Q9BM38
AC Q9BM38
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE LINE-LIKE REVERSE TRANSCRIPTASE (FRAGMENT).
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPONSON-LRT-L4 RETROTRANSPONSON;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.
RT "Transposable elements in sexual and asexual taxa."
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY013958; AAG59940.1; -
KW RNA-directed DNA polymerase.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2631 MW; 963CB9208057519A CRC64;

Query Match 17.5%; Score 35; DB 5; Length 25;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 12 SGPSLFNNFTVS 23
||||| :|||
Db 3 SSPSLFSPFTIN 14

RESULT 7
Q9S8NO PRELIMINARY; PRT; 29 AA.
ID Q9S8NO
AC Q9S8NO
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE P22 KUNITZ FAMILY OF TRYPSIN INHIBITOR (FRAGMENT).
OS Raphanus sativus (Radish).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Raphanus.
OX NCBI_TaxID=3726;
RN [1]
RP SEQUENCE.
RX MEDLINE=94272301; PubMed=8003938;
RA Lopez F., Vansuyt G., Derancourt J., Fourcroy P., Casse-Delbart F.;
RT "Identification by 2D-page analysis of salt-stress induced proteins in

RT radish (Raphanus sativus).";
RL Cell. Mol. Biol. 40:85-90(1994).
SQ SEQUENCE 29 AA; 3032 MW; F3937746B814F8B6 CRC64;

Query Match 17.5%; Score 35; DB 10; Length 29;
Best Local Similarity 35.7%; Pred. No. 4.2e+02;
Matches 10; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 4 SYGLRPGSSGSLFNNFTVSEWLRVPKV 31
||| :|||
Db 4 SYGTVTTPGPWL-----SAWLSAPQL 25

RESULT 8
O05602 PRELIMINARY; PRT; 32 AA.
ID O05602
AC O05602
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TRANSPONSON TN5041 DNA (FRAGMENT).
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KHP41; TRANSPONSON-TN5041;
RX MEDLINE=97419493; PubMed=9274008;
RA Kholidii G.Y., Yurieva O.V., Gorlenko Z.M., Mindlin S.Z., Bass I.A.,
RA Lomevskaya O.L., Kopteva A.V., Nikiforov V.G.;
RT "Tn5041 : a chimeric mercury resistance transposon closely related to
the toluene degradative transposon Tn4651.";
RL Microbiology 143:2549-2556(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-KHP41; TRANSPONSON-TN5041;
RA Kholidii G.Y., Mindlin S.Z., Gorlenko Z.M., Bass I.A., Kalyaeva E.S.,
RA Nikiforov V.;
RT "Host-dependent transposition of Tn5041.";
RL Russ. J. Genet. 36:365-373(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-KHP41; TRANSPONSON-TN5041;
RA Kholidii G.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; X98999; CAA67458.1; -
FT NON_TER 1 1
FT NON_TER 32 32
SQ SEQUENCE 32 AA; 3298 MW; AF42B5EEF917077A CRC64;

Query Match 17.5%; Score 35; DB 2; Length 32;
Best Local Similarity 46.7%; Pred. No. 4.6e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 10 GSSGSPSLFNNFTVSF 24
||| :|||
Db 1 GSCSPSLYCSFSSW 15

RESULT 9
Q9N1V7 PRELIMINARY; PRT; 32 AA.
ID Q9N1V7
AC Q9N1V7
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BETA-GLUCURONIDASE (FRAGMENT).
GN GUSB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]

```
RP SEQUENCE FROM N.A.
RX MEDLINE-20082971; PubMed-10613847;
RA Caetano A.R., Shiu Y.L., Lyons L.A., O'Brien S.J., Laughlin T.F.,
RT Bowling A.T., Murray J.D.;
RT "A comparative gene map of the horse (Equus caballus).";
RL Genome Res. 9:1239-1249(1999).
DR EMBL: AF134223; AAF63864.1; -.
DR HSSP: P08236; 1BHG.
FT NON_TER 1
FT NON_TER 32
SQ SEQUENCE 32 AA; 3805 MW; 5B55C4F8462720E9 CRC64;

Query Match 17.0%; Score 34; DB 6; Length 32;
Best Local Similarity 30.8%; Pred. No. 6.5e+02;
Matches 8; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

QY 10 GSSGSPSLFNFTVSVFWLRVPKVSASH 35
Db 5 GFDWPLLVDKDFNLLRWLGANSFRTSH 30

RESULT 10
O9XGG9 PRELIMINARY; PRT; 33 AA.
AC O9XGG9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE RIBONUCLEASE H (FRAGMENT).
GN RNASEH.
OS Equis sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-RPS10;
RA Pearce S.R., Stuart-Rogers C., Knox M., Kumar A., Ellis N.T.,
RA Flavell A.J.;
RT "Rapid isolation of plant Tyl-copia group retrotransposon LTR
sequences for molecular marker studies.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ243042; CAB4718.1; -.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3725 MW; C523ED4611598521 CRC64;

Query Match 17.0%; Score 34; DB 10; Length 33;
Best Local Similarity 42.1%; Pred. No. 6.7e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 16 LFNNFTVSVFWLRVPKVSAS 34
Db 12 LFDDFDSLNRPPVSVTT 30

RESULT 11
Q9UQB1 PRELIMINARY; PRT; 25 AA.
AC Q9UQB1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CD22 PROTEIN (FRAGMENT).
GN CD22.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99180618; PubMed-10079291;
RA Hatta Y., Tsuchiya N., Matsushita M., Shiota M., Hagiwara K.,
```

```
RA Tokunaga K.;
RT "Identification of the gene variations in human CD22.";
RL Immunogenetics 49:280-286(1999).
DR EMBL: AB013005; BAA36574.1; -.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 2969 MW; 354944648D2682B9 CRC64;

Query Match 16.5%; Score 33; DB 4; Length 25;
Best Local Similarity 58.3%; Pred. No. 6.9e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 GLRPGSSGSPSLF 17
Db 9 GLQENSSGSGSF 20

RESULT 12
Q95LC4 PRELIMINARY; PRT; 33 AA.
AC Q95LC4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE BETA-1 ADRENERGIC RECEPTOR (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21540585; PubMed-11683720;
RA Lee J.H., Zhang W., Moran C.;
RT "Comparative porcine gene mapping relative to human chromosomes 9, 10,
RT 20, and 22.";
RL Anim. Genet. 32:313-315(2001).
DR EMBL: AF345500; AAK57992.1; -.
KW Receptor.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3694 MW; 7DBEBDC3E268E150 CRC64;

Query Match 16.5%; Score 33; DB 6; Length 33;
Best Local Similarity 25.9%; Pred. No. 9.4e+02;
Matches 7; Conservative 1; Mismatches 7; Indels 12; Gaps 1;

QY 2 HWSYGLRPGSSGSPSLFNFTVSVFWLRV 28
Db 11 HWDYG-----SFFCELWTSV 25

RESULT 13
Q9A6D5 PRELIMINARY; PRT; 33 AA.
AC Q9A6D5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE HYPOTHETICAL PROTEIN CC2159.
GN CC2159.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocky I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khoufi H., Shetty J., Berry K.,
```

```

RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RT Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RL "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005888; AAK24130.1; -.
DR TIGR: CC2159; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 33 AA; 3620 MW; B4279FDE1F0F3929 CRC64;
    Query Match      16.5%; Score 33; DB 16; Length 33;
    Best Local Similarity 46.7%; Pred. No. 9.4e+02;
    Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 8 RFGSSGPSLNFNTV 22
Db 4 RIARSGPFLGRFAL 18

RESULT 14
Q9QVF6 PRELIMINARY; PRT; 20 AA.
AC Q9QVF6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CARBONIC ANHYDRASE IV, CA IV.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=92147688; PubMed=1737787;
RA Waheed A., Zhu X.L., Sly W.S.;
RT "Membrane-associated carbonic anhydrase from rat lung. Purification,
RT characterization, tissue distribution, and comparison with carbonic
RT anhydrase IVs of other mammals.";
RL J. Biol. Chem. 267:3308-3311(1992).
SQ SEQUENCE 20 AA; 2275 MW; 6CB2AC1CD44DEB65 CRC64;
    Query Match      16.2%; Score 32.5; DB 11; Length 20;
    Best Local Similarity 44.4%; Pred. No. 6.4e+02;
    Matches 8; Conservative 2; Mismatches 3; Indels 5; Gaps 2;
QY 2 HWSYGLR---PGS--SGP 14
Db 3 HWXYEQAKEPNXXSGP 20

RESULT 15
Q9L8K1 PRELIMINARY; PRT; 13 AA.
AC Q9L8K1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE VANS (FRAGMENT).
GN VANSB.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BM4281;
RX MEDLINE=20307504; PubMed=10846225;
RA Dahl K.H., Lundblad E.W., Roekenes T.P., Olsvik O., Sundsfjord A.;
RT "Genetic linkage of the vanB2 gene cluster to Tn5382 in vancomycin
RT resistant enterococci and characterization of two novel insertion
RT sequences.";
RL Microbiology 146:1469-1479(2000).
DR EMBL: AF201896; AAF73374.1; -.
DR NON_TER 1
FT 1

```

```

SQ SEQUENCE 13 AA; 1502 MW; 8D0E282189F9672A CRC64;
    Query Match      16.0%; Score 32; DB 2; Length 13;
    Best Local Similarity 54.5%; Pred. No. 4.8e+02;
    Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 22 VSFWLKRVPKVS 32
Db 1 VLFWLDLPPTS 11

Search completed: October 10, 2002, 22:33:04
Job time : 31 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 17:16:48 ; Search time 60 Seconds
(without alignments)
68.496 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHMSYGLRPGSGSLFNFTVSFWLRVPRKVSASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 310939

Minimum DB seq length: 0

Maximum DB seq length: 37

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	116	58.0	31	21	AA192653
2	113	56.5	36	22	AA192653
3	113	56.5	36	22	AA192653
4	112	56.0	21	12	AA11896
5	112	56.0	21	17	AA192653
6	112	56.0	21	17	AA192653
7	112	56.0	21	18	AA192653
8	112	56.0	21	19	AA192653
9	112	56.0	21	20	AA192653
10	112	56.0	21	20	AA192653
11	112	56.0	21	21	AA192653

12	112	56.0	21	21	AA192653	Tetanus toxin T ce
13	112	56.0	21	21	AA192653	Foreign epitope P3
14	112	56.0	21	21	AA192653	Amino acid sequenc
15	112	56.0	21	21	AA192653	CD4+ T cell epitop
16	112	56.0	21	22	AA11784	Clostridium tetani
17	112	56.0	21	22	AA192653	Amino acid sequenc
18	112	56.0	21	22	AA192653	Universal tetanus
19	112	56.0	21	22	AA192653	Tetanus Toxoid uni
20	112	56.0	21	22	AA192653	Tetanus Toxoid uni
21	112	56.0	21	22	AA192653	HER-2 B cell pepti
22	112	56.0	21	22	AA192653	Tetanus Toxoid TT9
23	112	56.0	21	22	AA192653	Tetanus Toxoid TT
24	112	56.0	21	22	AA192653	Tetanus Toxoid 947
25	112	56.0	31	21	AA192653	PSMpep011 - P30 in
26	112	56.0	31	21	AA192653	PSMpep012 - P30 in
27	112	56.0	32	15	AA192653	LHRH-containing im
28	112	56.0	32	15	AA192653	Amyloid beta/tetan
29	108	54.0	32	16	AA192653	IgE CH4 region con
30	107	53.5	22	15	AA192653	Helper T cell epit
31	107	53.5	22	16	AA192653	Tetanus toxin help
32	107	53.5	22	17	AA192653	Tetanus toxin help
33	107	53.5	22	17	AA192653	Pathogen derived T
34	107	53.5	22	21	AA192653	T helper cell (Th)
35	107	53.5	22	21	AA192653	Unidentified pepti
36	107	53.5	22	22	AA192653	Amino acid sequenc
37	107	53.5	34	16	AA192653	IgE CH4 region con
38	106	53.0	21	21	AA192653	Tetanus toxin (TTD
39	103	51.5	19	21	AA192653	HLA class II bindi
40	95.5	47.8	20	19	AA192653	Tetanus toxin frag
41	94.5	47.2	20	19	AA192653	Universal helper T
42	91	45.5	21	18	AA192653	T-cell stimulatory
43	87.5	43.8	22	17	AA192653	Carrier peptide fo
44	79	39.5	14	19	AA192653	Tetanus Toxoid pro
45	79	39.5	14	21	AA192653	MHC class II assoc

ALIGNMENTS

RESULT 1
AA192653
ID AA192653 standard; Peptide: 31 AA.
XX AC AA192653;
XX DT 10-AUG-2000 (first entry)
XX DE PSMpep010 - P30 inserted in hPSM insertion position 6.
XX KW Foreign epitope; P2; prostate specific membrane antigen; vaccination;
XX KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
XX KW prostate cancer; cell-associated peptide antigen.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 6..26
XX FT /label= P30
XX PN WO200020027-A2.
XX PD 13-APR-2000.
XX PF 05-OCT-1999; 99WO-DK00525.
XX PR 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX PA (MEBI-) M & E BIOTECH AS.
XX PI Stelnaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;

XX WPI; 2000-349917/30.
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 PS Example 1; Page 117; 220pp; English.
 XX
 CC AAY2650-55 are peptides designed which correspond to the P2 and P30
 CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
 CC amino acids in each end. The flanking amino acids correspond to the
 CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g.
 CC T cell proliferation assays, but also for ELISA or other in vitro
 CC assays. The claims detail a method for inducing immune responses against
 CC weakly immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, hPSM, The
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
 CC method comprises effecting simultaneous presentation by antigen
 CC producing cells (APCs) of the animals immune system of: (1) at least 1
 CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the
 CC respective PA and including at least one foreign T helper epitope (e.g.
 CC P2 and/or P30) are also claimed. The method is used to treat prostate,
 CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and
 CC Her2, respectively.
 XX
 SQ Sequence 31 AA;
 Query Match 58.0%; Score 116; DB 21; Length 31;
 Best Local Similarity 100.0%; Pred. No. 6e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 16 LFNNFTVSFWLRPKVSASHLE 37
 Db 5 LFNNFTVSFWLRPKVSASHLE 26
 RESULT 2
 AAG63662
 ID AAG63662 standard; peptide; 36 AA.
 AC AAG63662;
 XX
 XX 29-OCT-2001 (first entry)
 DT
 XX
 DE Peptide comprising 5 conjugation sites for a pseudopeptide.
 XX
 KW Pseudopeptide; immunomodulation; adjuvant; antigen presenting cell;
 KW macrophage; dendritic cell; vaccine; autoimmune disease.
 XX
 OS Synthetic.
 XX
 PN WO200146127-A1.
 XX
 PD 28-JUN-2001.
 XX
 PF 22-DEC-1999; 99WO-IB02038.
 XX
 PR 22-DEC-1999; 99WO-IB02038.
 XX
 PA (OMPH-) OM-PHARMA.
 XX
 PI Bauer J, Martin OR, Rodriguez S;
 XX
 DR WPI; 2001-502469/55.
 XX
 PT New amphiphilic acylated pseudopeptides having a functionalized
 PT auxiliary spacer, useful as immunomodulators, e.g. as adjuvants in
 PT vaccines -
 XX

PS Example 3; Page 61; 166pp; French.
 XX
 CC The specification describes N-Acylated pseudopeptides, which have
 CC a neutral or charged acidic group at one terminal and a functionalized
 CC auxiliary spacer at the other. The pseudopeptides show immunomodulatory
 CC and adjuvant action, based on activation of antigen presenting cells
 CC (e.g. macrophages or dendritic cells), induction of differentiation of
 CC dendritic cells, induction of cytokine production and induction of
 CC maturation of immunocompetent cell strains originating from hematopoietic
 CC and lymphoid organs. They reinforce humoral and cellular immunity. They
 CC can be grafted onto antigens (to modulate immune response) or onto drugs
 CC (to improve the therapeutic activity or targeting). The pseudopeptides
 CC are thus useful in human or veterinary medicine as immunizing or
 CC diagnostic agents. Typically, they are used as adjuvants together with
 CC (or covalently bonded to) antigens for vaccination against viral,
 CC parasitic/protozoal, microbial or fungal infections; incubated with blood
 CC cells ex vivo, to render the cells immunocompetent before reintroduction
 CC in vivo; or used in therapy of certain autoimmune diseases. The
 CC pseudopeptides are useful as carriers for antigens or other therapeutic
 CC agents due to their ability to form non-covalent bonds via the
 CC hydrophobic or hydrophilic auxiliary spacer. The present sequence
 CC represents a peptide, which has 5 possible conjugation sites for the
 CC pseudopeptides of the invention.
 XX
 SQ Sequence 36 AA;
 Query Match 56.5%; Score 113; DB 22; Length 36;
 Best Local Similarity 88.0%; Pred. No. 2e-09;
 Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 13 GPSLFNNFTVSFWLRPKVSASHLE 37
 Db 12 GITEFNFTVSFWLRPKVSASHLE 36
 RESULT 3
 AAG63515
 ID AAG63515 standard; peptide; 36 AA.
 AC AAG63515;
 XX
 XX 15-OCT-2001 (first entry)
 DT
 DE A peptide which may be conjugated to pseudopeptides.
 XX
 KW Pseudopeptide; immunomodulation; adjuvant; antigen presenting cell;
 KW macrophage; dendritic cell; cytokine production; immunocompetent cell;
 KW autoimmune disease.
 XX
 OS Synthetic.
 XX
 PN WO200146126-A1.
 XX
 PD 28-JUN-2001.
 XX
 PF 21-DEC-2000; 2000WO-FR03650.
 XX
 PR 22-DEC-1999; 99WO-IB02038.
 XX
 PA (OMPH-) OM-PHARMA.
 XX
 PI Bauer J, Martin OR, Rodriguez S;
 XX
 DR WPI; 2001-496651/54.
 XX
 PT New amphiphilic acylated pseudopeptides having a functionalized
 PT auxiliary spacer, useful as immunomodulators, e.g. as adjuvants in
 PT vaccines -
 XX
 PS Example 3.3; Page 88; 267pp; French.
 XX
 CC The specification describes N-Acylated pseudopeptides, which have
 CC a neutral or charged acidic group at one terminal and a functionalized

CC auxiliary spacer at the other. The pseudopeptides show immunomodulatory
 CC and adjuvant action, based on activation of antigen presenting cells
 CC (e.g. macrophages or dendritic cells), induction of differentiation of
 CC dendritic cells, induction of cytokine production and induction of
 CC maturation of immunocompetent cell strains originating from hematopoietic
 CC and lymphoid organs. They reinforce humoral and cellular immunity. They
 CC can be grafted onto antigens (to modulate immune response) or onto
 CC drugs (to improve the therapeutic activity or targeting). The
 CC pseudopeptides are thus useful in human or veterinary medicine as
 CC immunizing or diagnostic agents. Typically, the pseudopeptides are used
 CC as adjuvants together with (or covalently bonded to) antigens for
 CC vaccination against viral, parasitic/protozoal, microbial or fungal
 CC infections; incubated with blood cells *ex vivo*, to render the cells
 CC immunocompetent before reintroduction *in vivo*; or used in therapy of
 CC certain autoimmune diseases. The present sequence represents a
 CC peptide which may be conjugated to pseudopeptides of the invention.

XX Sequence 36 AA;

Query Match 56.5%; Score 113; DB 22; Length 36;
 Best Local Similarity 88.0%; Pred. No. 2e-09;
 Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 GPSLENNFTVSFWRVPRKVSASHLE 37
 | : ||||| ||||| ||||| |||||
 Db 12 GITEFNFTVSFWRVPRKVSASHLE 36

RESULT 4

AAR11896
 ID AAR11896 standard; peptide; 21 AA.

XX AC AAR11896;

DT 19-JUL-1991 (first entry)

DE Immunogenic conjugate constituent peptide, TT3.

KW Malaria vaccine; major histocompatibility complex.

XX OS Tetanus toxin.

FH Key Location/Qualifiers

FT Peptide 1..14

FT /label= active fragment (claimed)

XX EP427347-A.

XX PD 15-MAY-1991.

XX PF 07-NOV-1990; 90EP-0202948.

XX PR 10-NOV-1989; 89IT-0022355.

XX PA (ENIE) ENIRICERCH SPA.

XX PI Bianchi E, Pessi A, Corradin G;

XX WPI; 1991-141874/20.

PT Synthetic peptide(s) used as universal carriers - for preparing
 PT immunogenic conjugates used as vaccines against Plasmodium
 PT falciparum

XX Claim 1; page 13; 16pp; English.

XX This peptide corresponds to residues 947-967 of Tetanus toxin. It can
 CC be used as a universal carrier for the prepn. of an immunogenic
 CC conjugate. It is covalently bound to a peptide or polysaccharide
 CC hapten derived from a pathogen. This conjugate can be used as
 CC a vaccine for malaria. This peptide is recognised by different T-
 CC helper cell clones in association with alleles of the human MHC.
 CC It contains 2 epitopes: (a) 953-967, recognised by DR5-restricted

CC clones; and (b) 947-960, recognised by all other DR and DP-
 CC restricted clones.

XX Sequence 21 AA;

Query Match 56.0%; Score 112; DB 12; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWRVPRKVSASHLE 37
 | : ||||| ||||| ||||| |||||
 Db 1 FNNFTVSFWRVPRKVSASHLE 21

RESULT 5

AAW06130
 ID AAW06130 standard; Peptide; 21 AA.

XX AC AAW06130;

DT 07-FEB-1997 (first entry)

DE Tetanus toxoid protein T-cell epitope.

XX KW Cholesteryl ester transfer protein; CEST; antigen; vaccine;
 KW cardiovascular disease; atherosclerosis; tetanus toxoid;
 KW T-cell epitope.

XX OS Clostridium tetani.

XX PN WO9634888-A1.

XX PD 07-NOV-1996.

XX PF 01-MAY-1996; 96WO-US06147.

XX PR 01-MAY-1995; 95US-0432483.

XX PA (TCEL-) T CELL SCI INC.

XX PI Rittershaus CW, Thomas LJ;

XX WPI; 1996-506103/50.

XX Cholesteryl ester transfer protein B cell epitope linked to T cell
 PT epitope - used to generate vaccine to regulate CEST activity for
 PT decreasing the risk of developing a cardiovascular disease e.g.
 PT atherosclerosis

XX Claim 11; Page 43; 72pp; English.

XX A helper T-cell epitope (AAW06130) comprises amino acids 947-967
 CC of tetanus toxoid protein. It can be utilised in novel peptide
 CC vaccines (see also AAW06129, AAW06132) also including B-cell
 CC epitope(s) from human or rabbit cholesteryl ester transfer
 CC protein (CEST) to elicit an immune response against endogenous
 CC CEST activity, thereby treating or preventing a cardiovascular
 CC disease, such as atherosclerosis.

XX Sequence 21 AA;

Query Match 56.0%; Score 112; DB 17; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWRVPRKVSASHLE 37
 | : ||||| ||||| ||||| |||||
 Db 1 FNNFTVSFWRVPRKVSASHLE 21

RESULT 6

AAR88397

ID AAR88397 standard; Peptide; 21 AA.

```
XX AAR88397;
XX 12-JUN-1996 (first entry)
XX DE T-cell antigen TT3 peptide.
XX KW T-antigen; vaccine; antibody; T-cell; T-lymphocyte;
XX KW alpha-helix; coiled-coil heterodimer; core peptide; subunit.
XX OS Synthetic.
XX PN W09531480-A1.
XX PD 23-NOV-1995.
XX PF 18-MAY-1995; 95WO-CA00293.
XX PR 18-MAY-1994; 94US-0245507.
XX PA (SPIS-) SPI SYNTHETIC PEPTIDES INC.
XX PI Cachia PJ, Hodges RS, Houston ME, Irvin RT, Kay CM;
XX PI Zhou NE;
XX DR WPI; 1996-010880/01.
XX Hetero:dimeric polypeptide immunogen in coiled-coil configuration
XX PT with different antigens on each sub:unit - useful in vaccines and
XX PT for antibody prodn..
XX PS Claim 7; Page 62; 95pp; English.
XX CC This T-cell antigen TT3 peptide may be attached to a core peptide
XX CC contained in one of the 2 subunits of an alpha-helical coiled-coil
XX CC heterodimer. Each core peptide is comprised of terminal and
XX CC internal AA repeat sequences. This peptide antigen is attached
XX CC to the core peptide through covalent linkages to certain AA of the
XX CC in a stable alpha-helical coiled-coil configuration having a 1:1
XX CC stoichiometry, and the peptide antigen is disposed toward the outer
XX CC surfaces of the configuration. The heterodimer may be used as a
XX CC synthetic vaccine (optionally multivalent) or to generate
XX CC antibodies.
XX SQ Sequence 21 AA;
Query Match 56.0%; Score 112; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 FNNFTVSFWLRVPKVSASHLE 37
DB 1 FNNFTVSFWLRVPKVSASHLE 21
RESULT 7
AAW46449
ID AAW46449 standard; Peptide; 21 AA.
XX AC AAW46449;
XX DT 18-MAY-1998 (first entry)
XX DE Broad range helper T cell epitope from the tetanus toxoid protein.
XX KW Cholesteryl ester transfer protein; CERP; cholesteryl ester;
XX KW high density lipoprotein; HDL; very low density lipoprotein; VLDL;
XX KW low density lipoprotein; LDL; T cell epitope; antibody;
XX KW DNA plasmid-based vaccine; broad range helper T cell epitope;
XX KW treatment; cardiovascular disease.
XX OS Clostridium tetani.
```

```
XX WO9741227-A1.
XX PD 06-NOV-1997.
XX PF 01-MAY-1997; 97WO-US07294.
XX PR 21-FEB-1997; 97US-0802967.
XX PR 01-MAY-1996; 96US-0640713.
XX PA (TCEL-) T CELL SCI INC.
XX PI Thomas LJ;
XX PI WPI; 1997-549731/50.
XX DR DNA plasmid-based vaccine encodes CERP B cell and helper T cell
XX PT epitope(s) - used for elevating high density lipoprotein levels, and
XX PT for treating cardiovascular disease
XX PS Disclosure; Page 44; 67pp; English.
XX CC The present sequence represents a broad range helper T cell epitope
XX CC of the tetanus protein. It can be used in DNA plasmid-based vaccines
XX CC against cholesteryl ester transfer proteins (CERPs). CERPs mediate the
XX CC transfer of cholesteryl esters from high density lipoprotein (HDL)
XX CC to very low density lipoprotein (VLDL) and low density lipoprotein (LDL),
XX CC and vice versa. An increased CERP activity produces an atherogenic
XX CC lipoprotein profile and induces atherosclerosis. A DNA plasmid-based
XX CC vaccine comprises sequences encoding at least one B cell epitope of CERP
XX CC linked in frame with at least one segment encoding a broad range helper
XX CC T cell epitope. The vaccines can be used to elevate the ratio of
XX CC circulating HDL to circulating LDL, VLDL or total cholesterol in a human.
XX CC It can also be used for decreasing the level of endogenous CERP activity
XX CC in a human. The vaccine can be used to produce anti-CERP antibodies in
XX CC vivo and for treating cardiovascular disease.
XX SQ Sequence 21 AA;
Query Match 56.0%; Score 112; DB 18; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 FNNFTVSFWLRVPKVSASHLE 37
DB 1 FNNFTVSFWLRVPKVSASHLE 21
RESULT 8
AAW67034
ID AAW67034 standard; peptide; 21 AA.
XX AC AAW67034;
XX DT 15-DEC-1998 (first entry)
XX DE Tetanus toxin fragment (residues 947-967).
XX KW Tetanus toxin; vaccine; antibody; carbohydrate peptide conjugate;
XX KW dendrimeric poly-lysine; epitope; tumour.
XX OS Clostridium tetani.
XX PN W09843677-A1.
XX PD 08-OCT-1998.
XX PF 27-MAR-1998; 98WO-EP01922.
XX PR 27-MAR-1997; 97US-0041726.
XX PA (INSP ) INST PASTEUR.
```

PI Bay S, Cantacuzene D, Leclerc C, Lo-man R;
XX WPI; 1998-557071/47.
XX Carbohydrate peptide conjugate used as vaccine - comprises carrier
PT with dendrimeric poly-lysine enabling multiple epitopes to be
PT covalently attached
XX
PS Disclosure; Page 13; 55pp; English.
XX The invention relates to a new carbohydrate peptide conjugate, which
CC comprises a carrier with a dendrimeric poly-lysine enabling multiple
CC epitopes to be covalently attached to it. Also claimed are: (1) an
CC antibody purified from biological fluid or cells of organisms
CC administered with the carbohydrate peptide conjugate, and (2) a diagnosis
CC kit comprising antigen-specific antibodies elicited by immunisation with
CC the carbohydrate peptide conjugate. The peptide conjugate, antibody and
CC diagnosis kit are used to provide pharmaceutical compositions and
CC vaccines against tumours. These can be used to support an immune response
CC against viral infections caused by hepatitis virus, HIV or cytomegalo
CC virus. They can be used to enhance immune responses, especially B- and T-
CC cell responses, of humans and animals against bacterial infections. The
CC carbohydrate peptide conjugate stimulates the antibody and T-cell
CC response without stimulating undesired immune responses. The composition
CC is capable of increasing the survival of tumour bearing humans and
CC animals. The present sequence corresponds to residues 947-967 of tetanus
CC toxin. The synthetic peptide corresponding to this sequence may be used
CC as an epitope in a carbohydrate peptide conjugate.
XX
SQ Sequence 21 AA;
Query Match 56.0%; Score 112; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 FNNFTVSFWLRVPKVSASHLE 37
DB 1 FNNFTVSFWLRVPKVSASHLE 21
RESULT 9
AAW67579
ID AAW67579 standard; peptide; 21 AA.
AC AAW67579;
XX
DT 02-MAR-1999 (first entry)
XX
DE T-cell epitope peptide #5 for chimeric fimbria/T-cell epitope peptide.
XX
KW Chimeric; non-typable Haemophilus influenzae; fimbria; T-cell epitope;
KW immunogenic composition; immune response.
XX
OS Synthetic.
XX
PN US5843464-A.
XX
PD 01-DEC-1998.
XX
PF 02-JUN-1995; 95US-0460502.
XX
PR 02-JUN-1995; 95US-0460502.
XX
PA (OHIS) UNIV OHIO STATE.
XX
PI Bakaletz LO, Kaumaya PTP;
XX
DR WPI; 1999-044514/04.
XX
PT Synthetic chimeric fimbria peptide - useful for vaccination against
PT non-typable Haemophilus influenzae
XX
PS Disclosure; Column 4; 16pp; English.

XX The invention relates to the manufacture of a synthetic chimeric peptide
CC comprising a non-typable Haemophilus influenzae fimbria peptide fused via
CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is
CC used in immunogenic compositions which induce an immune response against
CC non-typable Haemophilus influenzae. This sequence represents an example
CC of a T-cell epitope peptide used to generate the chimeric peptide.
XX
SQ Sequence 21 AA;
Query Match 56.0%; Score 112; DB 20; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 FNNFTVSFWLRVPKVSASHLE 37
DB 1 FNNFTVSFWLRVPKVSASHLE 21
RESULT 10
AAW73222
ID AAW73222 standard; Protein; 21 AA.
XX
AC AAW73222;
XX
DT 25-JAN-1999 (first entry)
XX
DE Tetanus toxoid epitope.
XX
KW Multispecific single chain antibody; antibody H22; tumour cell; therapy;
KW antibody-dependent cellular cytotoxicity; ADCC; HER 2/neu; infection;
XX epidermal growth factor receptor; breast cancer; ovarian cancer.
OS Synthetic.
XX
PN US5837243-A.
XX
PD 17-NOV-1998.
XX
PF 07-JUN-1996; 96US-0661052.
XX
PR 07-JUN-1996; 96US-0661052.
PR 07-JUN-1995; 95US-0484172.
XX
PA (MEDA-) MEDAREX INC.
XX
PI Deo YM, Goldstein J, Graziano R, Somasundaram C;
XX
DR WPI; 1999-023374/02.
XX
PT Specific killing of tumour cells - using a multi-specific molecule
PT comprising an anti-Fc receptor antibody and a portion which binds to
PT a target cell
XX
PS Example 7; Column 27; 57pp; English.
XX
CC This sequence represents a tetanus toxoid epitope and is recognised
CC by the multispecific single chain antibody designated H22. The
CC antibody can be used in the method of the invention for inducing
CC antibody-dependent cellular cytotoxicity (ADCC) against a tumour cell
CC which is characterised by overexpression of HER 2/neu or epidermal growth
CC factor receptor (EGFR), comprises contacting the tumour cell with a
CC multispecific protein molecule (preferably a single chain antibody)
CC comprising: (a) an anti-Fc receptor antibody or an antigen binding
CC fragment; (b) a portion which binds to HER 2/neu; and (c) a portion which
CC binds to EGFR. The method can be used for treating cancers especially
CC breast cancer or ovarian cancer. The multispecific antibody can also
CC be administered prophylactically to vaccinate a subject against infection
CC by a target cell.
XX
SQ Sequence 21 AA;
Query Match 56.0%; Score 112; DB 20; Length 21;

```
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 11
AAB45512
ID AAB45512 standard; Protein; 21 AA.
AC AAB45512;
XX
DT 26-FEB-2001 (first entry)
XX
DE Tetanus P30 epitope SEQ ID NO: 24.
XX
KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS Clostridium tetani.
XX
PN WO200065058-A1.
XX
PD 02-NOV-2000.
XX
PF 19-APR-2000; 2000WO-DK00205.
XX
PR 23-APR-1999; 99DK-0000552.
PR 06-MAY-1999; 99US-0132811.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Klysner S;
XX
DR WPI; 2000-672791/65.
XX
PT Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -
XX
XX
PS Example 6; Page 137; 172pp; English.
XX
CC The present invention is concerned with methods of treating asthma,
CC eosinophilla, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.
XX
XX
SQ Sequence 21 AA;
Query Match 56.0%; Score 112; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 12
AAY99876
ID AAY99876 standard; Protein; 21 AA.
XX
AC AAY99876;
XX
DT 03-OCT-2000 (first entry)
XX
XX

Best Local Similarity 100.0%; Score 112; DB 21; Length 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 13
AAY92626
ID AAY92626 standard; Protein; 21 AA.
XX
AC AAY92626;
XX
DT 10-AUG-2000 (first entry)
XX
DE Foreign epitope P30.
XX
KW Foreign epitope; P2; prostate specific membrane antigen; PSM; Her2;
KW Heregulin 2; Fibroblast growth factor 8b; FGF8b; vaccination;
KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW prostate cancer; cell-associated peptide antigen.
XX
OS Clostridium tetani.
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.

Tetanus toxin T cell epitope helper peptide P30.
Human; MAGE-10; tumour rejection antigen precursor; bladder cancer;
prostate cancer; lung cancer; cancer detection; oesophageal cancer;
head and neck cancer; melanoma; myeloma; sarcoma; immunogen;
tetanus toxin.
XX
OS Homo sapiens.
XX
PN WO200026407-A1.
XX
PD 11-MAY-2000.
XX
PF 15-OCT-1999; 99WO-US24258.
PR 30-OCT-1998; 98US-0183714.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Boon-Falleur T, Brasseur F, Rimoldi D, Deplaen E;
XX
DR WPI; 2000-451624/39.
XX
PT Determining presence of cancer in samples, especially useful for
PT detecting bladder, prostate and lung cancer comprises assaying sample
PT for expression of tumor rejection antigen precursor MAGE-10 -
XX
XX
PS Example 12; Page 14; 26pp; English.
XX
CC The present sequence is a tetanus toxin T cell epitope known as
CC Helper Peptide P30. Hybrids of this peptide and an immunogenic peptide
CC derived from tumour rejection antigen precursor MAGE-10 were used to
CC generate polyclonal antiserum against MAGE-10. MAGE-10 binding monoclonal
CC antibodies can be used to detect MAGE-10 expression. A correlation
CC between MAGE-10 expression and cancer has been discovered and thus by
CC determining the presence of MAGE-10, the presence of cancer can be
CC determined. MAGE-10 expression can be detected using an immunoassay, an
CC oligonucleotide hybridisation assay or via other standard techniques.
CC This method is especially useful for determining the presence of
CC bladder, oesophageal, head and neck, prostate or lung cancer, or
CC melanoma, myeloma or sarcoma.
XX
XX
SQ Sequence 21 AA;
Query Match 56.0%; Score 112; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 13
AAY92626
ID AAY92626 standard; Protein; 21 AA.
XX
AC AAY92626;
XX
DT 10-AUG-2000 (first entry)
XX
DE Foreign epitope P30.
XX
KW Foreign epitope; P2; prostate specific membrane antigen; PSM; Her2;
KW Heregulin 2; Fibroblast growth factor 8b; FGF8b; vaccination;
KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW prostate cancer; cell-associated peptide antigen.
XX
OS Clostridium tetani.
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
```

XX 05-OCT-1999; 99WO-DK00525.
 XX
 PR 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX
 PA (MEBI-) M & E BIOTECH AS.
 XX
 PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX
 DR WPI; 2000-349917/30.
 DR N-PSDB; AAA09461.
 XX
 PT Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 PS Example 1; Page 214; 220pp; English.
 XX
 CC The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, human
 CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
 CC simultaneous presentation by antigen producing cells (APCs) of the
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
 CC group derived from the PA and/or at least 1 B-cell group derived from the
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is
 CC foreign to the animal. Analogues of human PSM, human Her2 and
 CC human/murine FGF8b comprising a substantial part of all known and
 CC predicted CTL and B-cell epitopes of the respective PA and including at
 CC least one foreign T helper epitope (e.g. P2 and/or P30) are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer
 CC when the PA is human PSM, FGF8b and Her2, respectively.
 XX
 SQ Sequence 21 AA;
 Query Match 56.0%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 FNNFTVSFWLRVPKVSASHLE 37
 Db 1 FNNFTVSFWLRVPKVSASHLE 21
 RESULT 14
 AAY84428
 ID AAY84428 standard; Peptide; 21 AA.
 XX
 AC AAY84428;
 XX
 DT 25-JUL-2000 (first entry)
 XX
 DE Amino acid sequence of the tetanus toxoid P30 epitope.
 XX
 KW Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;
 KW tumour necrosis factor receptor; type II transmembrane protein;
 KW osteoclast differentiation; CSF-1; osteoclast activator;
 KW immune response; osteoporosis; bone resorption;
 KW tetanus toxoid P30 epitope.
 XX
 OS Clostridium tetani.
 XX
 PN WO200015807-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 13-SEP-1999; 99WO-DK00481.
 XX
 PR 15-SEP-1998; 98DK-0001164.
 PR 02-OCT-1998; 98US-0102896.
 XX

PA (MEBI-) M & E BIOTECH AS.
 XX
 PI Halkier T, Haaning J;
 XX
 DR WPI; 2000-271444/23.
 XX
 PT In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used
 PT to treat, prevent and ameliorate osteoporosis -
 XX
 PS Example; Page 106; 110pp; English.
 XX
 CC The present sequence represents the tetanus toxoid P30 epitope. It is
 CC used to create a fusion protein with murine osteoprotegerin ligand
 CC (OPGL). Osteoprotegerin is a secreted member of the tumour necrosis
 CC factor receptor family, which blocks osteoclastogenesis in a dose
 CC dependent manner. The OPGL protein is synthesised as a type II
 CC transmembrane protein. The murine and human OPGL polypeptides are 87%
 CC homologous. OPGL is a potent osteoclast differentiation factor when
 CC combined with CSF-1. It is not capable of inducing osteoclast
 CC differentiation in the absence of CSF-1. OPGL is also an activator of
 CC mature osteoclasts. The specification describes a method for the in vivo
 CC down-regulation of OPGL activity in an animal. The method comprises
 CC using at least one OPGL polypeptide or subsequence, and/or at least one
 CC OPGL analogue to induce an immune response in the animal. The method
 CC and OPGL polypeptide are useful for treating, preventing and ameliorating
 CC osteoporosis or other diseases or conditions characterised by excessive
 CC bone resorption.
 XX
 SQ Sequence 21 AA;
 Query Match 56.0%; Score 112; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 FNNFTVSFWLRVPKVSASHLE 37
 Db 1 FNNFTVSFWLRVPKVSASHLE 21
 RESULT 15
 AAY49260
 ID AAY49260 standard; peptide; 21 AA.
 XX
 AC AAY49260;
 XX
 DT 07-FEB-2000 (first entry)
 XX
 DE CD4+ T cell epitope P30TT fragment.
 XX
 KW Carrier protein; CD4+; T cell epitope; N6; N10; N19; immunogen;
 KW encapsulated bacteria.
 XX
 OS Synthetic.
 XX
 PN WO9955730-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO-IB00844.
 XX
 PR 27-APR-1998; 98GB-0008932.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Rappuoli R, Grandi G;
 XX
 DR WPI; 2000-023325/02.
 XX
 PT Carrier proteins containing CD4+ epitopes useful for protecting against
 PT diseases caused by encapsulated bacteria -
 XX
 PS Disclosure; Page 36; 76pp; English.
 XX

CC The invention provides carrier proteins comprising at least 5 CD4+ T
CC cell epitope. The carrier protein comprises at least 1 of N6, N10 or
CC N19. The carrier protein can be prepared by expressing a vector
CC comprising a nucleic acid molecule encoding the protein, in a host cell
CC and recovering the expressed protein. The carrier protein can also be
CC produced by (a) constructing oligonucleotide molecules that encode
CC peptide epitopes; (b) annealing the oligonucleotides to form duplexes;
CC (c) introducing the duplexes into an expression vector; (d) introducing
CC the expression vector into a host cell; and (e) isolating the fusion
CC protein produced from a culture of the host cells. The carrier protein
CC can be used as a protective immunogen in the control of diseases caused
CC by encapsulated bacteria. Sequences AA149256-266 represent CD4+ T cell
CC epitopes inserted in the recombinant polypeptide carrier proteins.

xx
SQ Sequence 21 AA;

Query Match 56.0%; Score 112; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRLVPKVSASHLE 37
| | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRLVPKVSASHLE 21

Search completed: October 10, 2002, 21:23:53
Job time : 61 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 21:22:43 ; Search time 24 Seconds
(without alignments)
37.656 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSSGSLFNNFTVFWLRPKVSASHLE 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 155891

Minimum DB seq length: 0
Maximum DB seq length: 37

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA: *
1: /cqn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cqn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cqn2_6/ptodata/1/1aa/6A_COMB.pep: *
4: /cqn2_6/ptodata/1/1aa/6B_COMB.pep: *
5: /cqn2_6/ptodata/1/1aa/PCTUS_COMB.pep: *
6: /cqn2_6/ptodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	57.0	31	5	PCT-US93-11703-64
2	112	56.0	21	1	US-07-610-525-1
3	112	56.0	21	2	US-08-661-052-12
4	112	56.0	21	2	US-08-460-502-8
5	112	56.0	21	2	US-08-724-7748-5
6	112	56.0	21	4	US-09-089-595-5
7	112	56.0	21	4	US-09-382-855-5
8	112	56.0	21	4	US-09-183-7148-5
9	112	56.0	21	4	US-09-188-082-12
10	112	56.0	21	4	US-09-171-969-10
11	112	56.0	21	5	PCT-US93-11703-66
12	112	56.0	32	1	US-08-446-692-14
13	112	56.0	32	2	US-08-488-351A-14
14	107	53.5	22	1	US-08-446-692-5
15	107	53.5	22	2	US-08-488-351A-5
16	107	53.5	22	3	US-09-100-409A-41
17	107	53.5	22	5	PCT-US95-13841-8
18	100	50.0	19	1	US-07-610-525-2
19	94.5	47.2	20	2	US-08-319-704-11
20	88	44.0	17	1	US-07-610-525-3
21	87.5	43.8	22	2	US-08-817-933A-8
22	79	39.5	14	1	US-07-610-525-6
23	79	39.5	14	1	US-08-787-547-43
24	79	39.5	15	1	US-07-610-525-4
25	79	39.5	15	1	US-08-787-547-44
26	71	35.5	17	1	US-08-188-223-6
27	71	35.5	17	4	US-08-968-466-6

28	71	35.5	17	4	US-08-478-546B-6	Sequence 6, Appli
29	69	34.5	13	1	US-07-610-525-5	Sequence 5, Appli
30	62	31.0	14	1	US-08-453-588-23	Sequence 23, Appli
31	62	31.0	14	3	US-08-521-079-23	Sequence 23, Appli
32	62	31.0	17	1	US-07-690-983D-18	Sequence 18, Appli
33	62	31.0	18	1	US-07-690-983D-28	Sequence 28, Appli
34	61	30.5	16	1	US-08-453-588-26	Sequence 26, Appli
35	61	30.5	16	1	US-08-188-223-2	Sequence 2, Appli
36	61	30.5	16	3	US-08-521-079-26	Sequence 26, Appli
37	61	30.5	16	4	US-08-968-466-2	Sequence 2, Appli
38	61	30.5	16	4	US-08-478-548B-2	Sequence 2, Appli
39	60	30.0	18	4	US-09-082-279B-1147	Sequence 1147, Ap
40	60	30.0	18	4	US-09-315-304B-1147	Sequence 1147, Ap
41	59	29.5	12	1	US-08-453-588-24	Sequence 24, Appli
42	59	29.5	12	3	US-08-521-079-24	Sequence 24, Appli
43	59	29.5	14	1	US-07-690-983D-22	Sequence 22, Appli
44	59	29.5	14	1	US-07-690-983D-24	Sequence 24, Appli
45	59	29.5	14	1	US-07-690-983D-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1
PCT-US93-11703-64
; Sequence 64, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-11703-64

Query Match 57.0%; Score 114; DB 5; Length 31;
Best Local Similarity 95.5%; Pred. No. 1.6e-10;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 16 LFNFTVSWLRPKVSASHLE 37
:|||||
Db 6 MFNFTVSWLRPKVSASHLE 27

```
RESULT 2
US-07-610-525-1
; Sequence 1, Application US/07610525
; Patent No. 5196512
; GENERAL INFORMATION:
; APPLICANT: BIANCHI Elisabetta
; APPLICANT: PESSI Antonello
; APPLICANT: CORRADIN Giampietro
; TITLE OF INVENTION: SYNTHETIC PEPTIDES USEFUL AS
; TITLE OF INVENTION: UNIVERSAL CARRIERS FOR THE PREPARATION OF IMMUNOGENIC CONJUGATES
; TITLE OF INVENTION: AND THEIR USE IN THE DEVELOPMENT OF SYNTHETIC VACCINES.
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHEA & GOULD
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10020-1193
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/610,525
; FILING DATE: 19901108
; CLASSIFICATION: 424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-827-3000
; TELEFAX: 212-840-6702
; TELEX: 423973
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acid residues
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOPHETICAL: no
; FRAGMENT TYPE: Internal fragment
;
US-07-610-525-1
Query Match 56.0%; Score 112; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRVPKVSASHLE 37
Db 1 FNNFTVSWLRVPKVSASHLE 21
|||||
RESULT 3
US-08-661-052-12
; Sequence 12, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezhian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER-READABLE FORM:
;
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/661,052
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/484,172
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MXI-043CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
US-08-661-052-12
Query Match 56.0%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRVPKVSASHLE 37
Db 1 FNNFTVSWLRVPKVSASHLE 21
|||||
RESULT 4
US-08-460-502-8
; Sequence 8, Application US/08460502
; Patent No. 5843464
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldrick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

; MOLECULE TYPE: peptide

US-08-460-502-8

Query Match 56.0%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
|||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 5

US-08-724-774B-5
; Sequence 5, Application US/08724774B
; Patent No. 5908778

; GENERAL INFORMATION:

; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
; APPLICANT: Pierref; Cerrottini, Jean-Charles; Carrel,
; APPLICANT: Stefan; Reed, Daryl

; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
; TITLE OF INVENTION: Rejection Antigen Precuros Mage-10,
; TITLE OF INVENTION: Antibodies Specific To The Molecule, and

; TITLE OF INVENTION: Uses Thereof

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,774B

; FILING DATE: 03-October-1996

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5908778man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5457

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21

; TYPE: amino acid

; TOPOLOGY: linear

US-08-724-774B-5

Query Match 56.0%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
|||||
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 6

US-09-089-595-5

; Sequence 5, Application US/09089595
; Patent No. 6153728

; GENERAL INFORMATION:

; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
; APPLICANT: Pierref; Cerrottini, Jean-Charles; Carrel,
; APPLICANT: Stefan; Reed, Daryl

; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor

; TITLE OF INVENTION: Rejection Antigen Precuros Mage-10,
; TITLE OF INVENTION: Antibodies Specific To The Molecule, and
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/089,595

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,774

; FILING DATE: 03-October-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 6153728man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 5457

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 21

; TYPE: amino acid

; TOPOLOGY: linear

US-09-089-595-5

Query Match 56.0%; Score 112; DB 4; Length 21;

Best Local Similarity 100.0%; Pred. No. 2e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37

|||||

Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 7

US-09-382-855-5

; Sequence 5, Application US/09382855

; Patent No. 6174692

; GENERAL INFORMATION:

; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie, Pierre;

; APPLICANT: Cerrottini, Jean-Charles; Carrel, Stefan; Reed, Daryl

; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor Rejection

; TITLE OF INVENTION: Antigen Precuros Mage-10, Antibodies Specific To The Molecu

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fulbright & Jaworski LLP

; STREET: 666 Fifth Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/382,855

; FILING DATE: 25-August-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 09/089,595

```
; FILING DATE: 02-June-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Schofield, Mary Anne
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
; US-09-382-855-5

Query Match 56.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 8
US-09-183-714B-5
; Sequence 5, Application US/09183714B
; Patent No. 6221593
; GENERAL INFORMATION:
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: Brasseur, Francis
; APPLICANT: Rimoldi, Donata
; APPLICANT: De Plaen, Etienne
; TITLE OF INVENTION: Method for Determining Cancer by Determining Expression
; FILE REFERENCE: of MAGE-10
; CURRENT APPLICATION NUMBER: US/09/183,714B
; CURRENT FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 08/724,774
; PRIOR FILING DATE: 1996-10-03
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; US-09-183-714B-5

Query Match 56.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 9
US-09-188-082-12
; Sequence 12, Application US/09188082
; Patent No. 6270765
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,082
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/661,052
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-09-188-082-12

Query Match 56.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10
US-09-171-969-10
; Sequence 10, Application US/09171969
; Patent No. 6284533
; GENERAL INFORMATION:
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: PLASMID-BASED VACCINE FOR TREATING ATHEROSCLEROSIS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 75 State Street, Suite 2300
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1807
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,969
; FILING DATE: 01 May 1997 (01.05.97)
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/640,713
; FILING DATE: 01 May 1996 (01.05.96)
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/802,967
; FILING DATE: 21 February 1997 (21.02.97)
```

ATTORNEY/AGENT INFORMATION:
NAME: Leon R. Yankwich
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FEATURE:
NAME/KEY:
LOCATION:
US-09-171-969-10

Query Match 56.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRVPKVSASHLE 37
Db 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 11
PCT-US93-11703-66
SEQUENCE 66, Application PC/TUS9311703
GENERAL INFORMATION:
APPLICANT: Chiron Mimotopes Pty. Ltd.
TITLE OF INVENTION: T-Cell Epitopes
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grant D. Green
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11703
FILING DATE: 28-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,852
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0222.101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-11703-66

Query Match 56.0%; Score 112; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRVPKVSASHLE 37

Db 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 12
US-08-446-692-14
SEQUENCE 14, Application US/08446692
PATENT NO. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-14

Query Match 56.0%; Score 112; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRVPKVSASHLE 37
Db 3 FNNFTVSWLRVPKVSASHLE 23

RESULT 13
US-08-488-351A-14
SEQUENCE 14, Application US/08488351A
PATENT NO. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 14:
LENGTH: 32 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-14

Query Match 56.0%; Score 112; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 3 FNNFTVSFWLRVPKVSASHLE 23

RESULT 14
US-08-446-692-5
Sequence 5, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-5

Query Match 53.5%; Score 107; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHL 36
Db 3 FNNFTVSFWLRVPKVSASHL 22

RESULT 15
US-08-488-351A-5
Sequence 5, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-5

Query Match      53.5%; Score 107; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 FNNFTVSEFLRVPKVSASHL 36
    |||
Db 3 FNNFTVSEFLRVPKVSASHL 22
    |||

```

Search completed: October 10, 2002, 22:35:36
Job time : 24 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:45:57 ; Search time 12.5 seconds
(without alignments)
238.302 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSYGLRPGSSGSLQYKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 6607

Minimum DB seq length: 0
Maximum DB seq length: 31

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	35.8	10	1 RHPGG	gonadoliberin - pi
2	58	35.8	10	1 RSHG	gonadoliberin - sh
3	54	33.3	10	1 RRAQ1	gonadoliberin I -
4	45	27.8	10	2 A2114	gonadoliberin - ch
5	40	24.7	10	1 RRAQ2	gonadoliberin II -
6	40	24.7	10	1 A61126	gonadoliberin - sp
7	40	24.7	10	2 A46030	gonadoliberin I -
8	40	24.7	10	2 B46030	gonadoliberin II -
9	37	22.8	10	2 A49187	gonadotropin-relea
10	36	22.2	26	2 S00579	hypothetical prote
11	36	22.2	31	2 PN0050	unidentified QM002
12	33	20.4	30	2 I49412	gamma-crystallin-3
13	31	19.1	21	2 C39543	collagen alpha 3(I
14	30.5	18.8	26	2 S05414	steryl-sulfatase (
15	30.5	18.8	27	2 S22355	hypothetical prote
16	30	18.5	24	2 PH0084	fatty acid-binding
17	30	18.5	24	2 S40139	T-cell receptor J-
18	30	18.5	27	2 PN0584	tyrosine 3-monooxy
19	30	18.5	30	2 S72626	small-cell-variant
20	29	17.9	18	2 S09723	2S albumin small c
21	29	17.9	20	2 A39543	collagen alpha 1(I
22	29	17.9	24	2 T50123	peroxisomal target
23	29	17.9	26	2 S78761	ribosomal protein
24	29	17.9	27	2 PN0583	tyrosine 3-monooxy
25	29	17.9	27	2 PN0585	tyrosine 3-monooxy
26	29	17.9	27	2 PN0586	tyrosine 3-monooxy
27	29	17.9	30	2 A49955	protein-tyrosine k
28	29	17.9	31	2 A70219	hypothetical prote
29	28.5	17.6	22	2 PQ0070	T-cell receptor be

30	28	17.3	21	2 JU0386	nitrile hydratase
31	28	17.3	26	2 S58387	T-cell receptor al
32	27	16.7	12	2 S70337	napiin small chain
33	27	16.7	12	2 S43170	kinesin light chai
34	27	16.7	15	2 F48223	synaptosomal-assoc
35	27	16.7	18	2 S09722	2S albumin small c
36	27	16.7	23	2 I50174	collagen alpha 1(I
37	27	16.7	27	2 S05219	photosystem I chai
38	26.5	16.4	27	1 T2A23	toxin III - snake-
39	26.5	16.4	29	2 A60604	glutathione peroxi
40	26	16.0	7	2 A44428	platelet aggregati
41	26	16.0	15	2 B41436	ovostatin - green
42	26	16.0	16	2 JT0609	leukocyte chemoatt
43	26	16.0	17	4 I51887	hypothetical EMSK
44	26	16.0	20	2 PQ0003	chlorophyll a/b-bi
45	26	16.0	20	2 A38600	trypsin inhibitor

ALIGNMENTS

RESULT 1

RHPGG

gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411

R:baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90172; MUID:72114303

A:Accession: A01411

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A>Note: the synthetic and natural hormones have the same physicochemical and biologic

R:baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A>Note: Trp-3 appears to be essential for biological activity

C:Comment: this hypothalamic hormone stimulates the secretion of both luteinizing and

C:Superfamily: gonadoliberin

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.8%; Score 58; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.027;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10

RESULT 2

RSHG

gonadoliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

A:Residues: 1-10 <BUR>
A:Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.027; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 3
RHAQ1
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Lance, V.; Swanson, R.; Regal, P.E.; 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.1; 0; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 4
A21114
gonadoliberin - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match 27.8%; Score 45; DB 2; Length 10;
Best Local Similarity 77.8%; Pred. No. 2.1; 2; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 5
RHAQ2
gonadoliberin II - American alligator
N:Alternate names: gonadotropin-releasing hormone II

C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: B60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Lance, V.; Swanson, R.; Regal, P.E.; 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of American alligator
A:Reference number: A60066; MUID:91352338
A:Accession: B60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 24.7%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 1; 1; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 6
A61126
gonadoliberin - spotted ratfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Hydroloagus colliel (spotted ratfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C:Accession: A61126
R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.
Gen. Comp. Endocrinol. 82, 152-161, 1991
A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a holostean fish
A:Reference number: A61126; MUID:91340067
A:Accession: A61126
A:Molecule type: protein
A:Residues: 1-10 <LOV>
A:Experimental source: brain
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 24.7%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 1; 1; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 7
A46030
gonadoliberin I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, P.; 33, 105-116, 1991
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 24.7%; Score 40; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 1; 1; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 |||:| ||
 Db 2 HWSHGWLP 10

RESULT 8

B46030
 gonadoliberin II - spiny dogfish
 N:Alternate names: gonadotropin-releasing hormone
 C:Species: Squalus acanthias (spiny dogfish)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
 C:Accession: B46030
 R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
 Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
 A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
 A:Reference number: A46030; MUID:92335300
 A:Accession: B46030
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <LOV>
 C:Keywords: hormone; prolutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 24.7%; Score 40; DB 2; Length 10;

Best Local Similarity 66.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 |||:| ||
 Db 2 HWSHGWTP 10

RESULT 9

A49187
 gonadotropin-releasing hormone III - sea lamprey
 C:Species: Petromyzon marinus (sea lamprey)
 C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
 C:Accession: A49187
 R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
 Endocrinology 132, 1125-1131, 1993
 A:Title: Primary structure and biological activity of a third gonadotropin-releasing hor
 A:Reference number: A49187; MUID:93178316
 A:Accession: A49187
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <SOW>
 A:Experimental source: brain
 A>Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 22.8%; Score 37; DB 2; Length 10;

Best Local Similarity 55.6%; Pred. No. 30;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 |||:| ||
 Db 2 HWSHGWPG 10

RESULT 10

S00579
 hypothetical protein 547 (hisa 5' region) - Methanococcus vannielii (fragment)
 C:Species: Methanococcus vannielii
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 22-Oct-1999
 C:Accession: S00579
 R:Brown, J.W.; Thomm, M.; Beckler, G.S.; Frey, G.; Stetter, K.O.; Reeve, J.N.
 Nucleic Acids Res. 16, 135-150, 1988
 A:Title: An archaeobacterial RNA polymerase binding site and transcription initiation of
 A:Reference number: S00579; MUID:88124191
 A:Accession: S00579
 A>Status: translation not shown
 A:Molecule type: DNA

A:Residues: 1-26 <BRO>
 A:Cross-references: EMBL:X07391; NID:g44714; PIDN:CAA30299.1; PID:g44715

Query Match 22.2%; Score 36; DB 2; Length 26;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 10 GSSGSPSLQYIKANSK 24
 | | | | | | | | | |
 Db 7 GHSGLVLQYIKEVEK 21

RESULT 11

PN0050
 unidentified QM0026 protein - mouse (fragments)
 C:Species: Mus musculus (house mouse)
 C:Date: 29-Oct-1997 #sequence_revision 23-Jan-1998 #text_change 23-Jan-1998
 C:Accession: PN0050
 R:Kato, H.
 Kawasaki Igakkaishi 22, 245-259, 1996
 A:Title: Analysis of proteins isolated by two dimensional electrophoresis of mouse n
 A:Reference number: PN0041
 A:Accession: PN0050
 A:Molecule type: protein
 A:Residues: 1-31 <KAT>
 A:Experimental source: neuroblastoma cell
 C:Comment: The molecular mass is 28,300 and the pI is 6.25. The amino-terminus is blo
 C:Keywords: brain

Query Match 22.2%; Score 36; DB 2; Length 31;

Best Local Similarity 25.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 6 GLRPGSSPSLQYIKANSKFEGIT 29
 | | | | | | | | | |
 Db 3 GLXLXGDEAPNXYLYRTPQLIXLS 26

RESULT 12

I49412
 gamma-crystallin-3 - western wild mouse (fragment)
 C:Species: Mus spretus (western wild mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
 C:Accession: I49412
 R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau,
 Mamm. Genome 5, 349-355, 1994
 A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
 A:Reference number: I48934; MUID:94319082
 A:Accession: I49412
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-30 <RES>
 A:Cross-references: EMBL:U05704; NID:g497029; PIDN:AAB60467.1; PID:g497030
 C:Superfamily: beta-crystallin
 C:Keywords: duplication

Query Match 20.4%; Score 33; DB 2; Length 30;

Best Local Similarity 45.0%; Pred. No. 3.9e+02;
 Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 5 YGLRPGSSGSPSLQYIKANSK 24
 | | | | | | | | | |
 Db 6 YLLRPGMYRRYLDWGAAAK 25

RESULT 13

C39543
 collagen alpha 3(IX) chain - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 23-May-1997
 C:Accession: C39543
 R:Wu, J.J.; Lark, M.W.; Chun, L.E.; Eyre, D.R.
 J. Biol. Chem. 266, 5625-5628, 1991

A;Title: Sites of stromelysin cleavage in collagen types II, IX, X, and XI of cartilage.
A;Reference number: A39543; MUID:91170231
A;Accession: C39543
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-21 <WUA>
C;Superfamily: unassigned collagens

Query Match 19.1%; Score 31; DB 2; Length 21;
Best Local Similarity 71.4%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 RPOSSGP 14
| | | : | |
Db 12 RGPAGP 18

RESULT 14

S05414
steryl-sulfatase (EC 3.1.6.2), microsomal - rat (fragment)
N;Alternate names: arylsulfatase C
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 28-Apr-1993
C;Accession: S05414
R;Rawano, J.I.; Kotani, T.; Ohtaki, S.; Minamino, N.; Matsuo, H.; Oinuma, T.; Aikawa, E.
Biochim. Biophys. Acta 997, 199-205, 1989
A;Title: Characterization of rat and human steroid sulfatases.
A;Reference number: S05414; MUID:89352671
A;Accession: S05414
A;Molecule type: protein
A;Residues: 1-26 <KAW>
C;Keywords: sulfuric ester hydrolase

Query Match 18.8%; Score 30.5; DB 2; Length 26;
Best Local Similarity 45.0%; Pred. No. 7.7e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 8 RPOSSGPSQYIKANSKFIG 27
| | | | | : | : | |
Db 2 RPG-PGNFLLINADDLGIG 20

RESULT 15

S52355
hypothetical protein - human
C;Species: Homo sapiens (man)
C;Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
C;Accession: S52355
R;Agrykastaritis, A.; Leversha, M.A.; Ferguson-Smith, M.; Moschonas, M.K.
submitted to the EMBL Data Library, March 1993
A;Description: A cosmid clone mapped to human chromosome 11p15 detects a Tag I restriction
A;Reference number: S52355
A;Accession: S52355
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-27 <AGR>
A;Cross-references: EMBL:X72878; NID:g666997; PIDN:CAA51390.1; PID:g666998

Query Match 18.8%; Score 30.5; DB 2; Length 27;
Best Local Similarity 38.1%; Pred. No. 8.1e+02;
Matches 8; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

QY 4 SYGLRPGSSGPSQYIKANSK 24
| | | | | : | : | |
Db 3 SLGXEG-GGKMIFYLRTNSK 22

Search completed: October 10, 2002, 16:49:59
Job time : 12.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:41:22 ; Search time 9.5 Seconds
(without alignments)
126.348 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSYGLRPGSSGPSQYIKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 1969

Minimum DB seq length: 0
Maximum DB seq length: 31

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	54	33.3	10	1	GONL_ALLMI	P37041 alligator m
2	47	29.0	10	1	GONL_CLUPA	P81749 clupea pall
3	45	27.8	10	1	GON3_ONCRH	P20367 oncorhynch
4	40	24.7	10	1	GON3_CHICK	P37043 gallus gall
5	40	24.7	10	1	GONL_SQUAC	P27429 squalus aca
6	37	22.8	10	1	GON3_PETMA	P30948 petromyzon
7	32	19.8	10	1	GONL_CHEPR	P80677 chelyosoma
8	28	17.3	10	1	GON2_CHEPR	P80678 chelyosoma
9	28	17.3	21	1	NRLA_ACISP	P33036 acinetobact
10	27	16.7	27	1	PSAF_SYNVU	P23078 synecchococ
11	26.5	16.4	19	1	TRP3_LEUMA	P81735 leucophaea
12	26.5	16.4	27	1	TXA3_ANESU	P01535 anemonia su
13	26	16.0	27	1	LS2_ADE07	P05663 human adeno
14	26	16.0	28	1	GUN_SCHCO	P81190 schizophyll
15	26	16.0	31	1	PETL_OENHO	Q9mtk4 oenothera h
16	25	15.4	10	1	GRP_RANRI	P33260 rana ridibu
17	25	15.4	17	1	UC06_MAIZE	P80612 zea mays (m
18	25	15.4	15	1	RL24_PROVU	P20032 proteus vul
19	25	15.4	20	1	NLH1_HELAN	P82007 helianthus
20	25	15.4	21	1	RL21_HALCU	P05974 halobacteri
21	25	15.4	22	1	CR33_LITCE	P56240 litoria cae
22	25	15.4	22	1	CR34_LITCE	P56241 litoria cae
23	25	15.4	23	1	GRP_ONCMY	O9ps30 oncorhynch
24	25	15.4	25	1	GRP_SCYCA	P09472 scyllorhinu
25	25	15.4	27	1	GRP_CANFA	P08989 canis fami
26	25	15.4	27	1	GRP_CHICK	P01295 gallus gall
27	25	15.4	27	1	GRP_PIG	P01294 sus scrofa
28	25	15.4	28	1	GRP_ALLMI	P31886 alligator m
29	25	15.4	29	1	PSAF_SYNP6	P31083 synecchococ
30	24.5	15.1	30	1	FIBR_PANIN	P22775 panulirus i
31	24	14.8	8	1	ALU1_CYPDPO	P82152 cydia pomon
32	24	14.8	10	1	GONL_PETMA	P04378 petromyzon
33	24	14.8	11	1	TKC2_CALVO	P41518 calliphora

34	24	14.8	13	1	TEML_RANTE	P57104 rana tempor
35	24	14.8	18	1	AGI_EUPMA	P33889 euphorbia m
36	24	14.8	18	1	RL24_SERMA	P49624 serratia ma
37	24	14.8	18	1	SODM_MYCHA	P80582 mycobacteri
38	24	14.8	20	1	AMP_FUSNU	P81207 fusobacteri
39	24	14.8	20	1	CRP_MUSCA	P19094 mustelus ca
40	24	14.8	25	1	COXO_ONCMY	P80334 oncorhynch
41	24	14.8	28	1	P222_MICNI	P21791 micrurus ni
42	24	14.8	30	1	KAB5_OLDAF	P58456 oldenlandia
43	24	14.8	30	1	Y161_TREPA	O83196 treponema p
44	23.5	14.5	26	1	PCW4_PACGO	P82426 pachycondyl
45	23	14.2	8	1	AL16_CARMA	P81819 carcinus ma

ALIGNMENTS

RESULT 1						
GONL_ALLMI						
ID	GONL_ALLMI	STANDARD;	PRT;	10	AA.	
AC	P37041; P20407;					
DT	01-FEB-1991 (Rel. 17, Created)					
DT	01-FEB-1991 (Rel. 17, Last sequence update)					
DT	15-DEC-1998 (Rel. 37, Last annotation update)					
DE	Gonadoliberein I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)					
DE	(Luliberin I).					
OS	Alligator mississippiensis (American alligator).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Archosauria; Crocodylidae; Alligatorinae; Alligator.					
OX	NCBI_TaxID=8496;					
RN	[1]					
RP	SEQUENCE.					
RC	TISSUE=Brain;					
RX	MEDLINE=91352338; PubMed=1882082;					
RA	Lovely D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,					
RA	Lance V., Swanson P., Rivier J.E., Sherwood N.M.;					
RT	"Primary structure of two forms of gonadotropin-releasing hormone					
RT	from brains of the American alligator (Alligator mississippiensis).";					
RL	Regul. Pept. 33:105-116(1991).					
CC	-!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.					
CC	-!- SUBCELLULAR LOCATION: Secreted.					
CC	-!- SIMILARITY: BELONGS TO THE GNRH FAMILY.					
DR	PIR; A60066; RHAQ1.					
DR	InterPro; IPR002012; GnRH.					
DR	Pfam; PF00446; GnRH; 1.					
DR	PROSITE; PS00473; GnRH; 1.					
DR	Hormone; Amidation; Hypothalamus.					
FT	MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.					
FT	MOD_RES 10 10 AMIDATION.					
SQ	SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;					

Query Match 33.3%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.025; 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 0; Indels 0;

QY 2 HWSYGLRPG 10
DB 2 HWSYGLQPG 10
|||||:|

RESULT 2						
GONL_CLUPA						
ID	GONL_CLUPA	STANDARD;	PRT;	10	AA.	
AC	P81749;					
DT	16-OCT-2001 (Rel. 40, Created)					
DT	16-OCT-2001 (Rel. 40, Last sequence update)					
DT	16-OCT-2001 (Rel. 40, Last annotation update)					
DE	Gonadoliberein I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH)					
DE	(Luliberin I).					
GN	GNRH1.					
OS	Clupea pallasii (Pacific herring).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;					

OC Clupeinae; Clupea.
 OX NCBI_TaxID=30724;
 RN [1]
 RP SEQUENCE, AND FUNCTION
 RC TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
 RA Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring."
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;
 Query Match 29.0%; Score 47; DB 1; Length 10;
 Best Local Similarity 77.8%; Pred. No. 0.28;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 III:II II
 DB 2 HWSHGLSPG 10
 III:II II
 RESULT 3
 GON3_ONCKE STANDARD; PRT; 10 AA.
 AC P20367; P81751;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III) (LH-RH III) (Luliberin III).
 GN GNRH3
 OS Oncorhynchus keta (Chum salmon), and
 OC Clupea pallasi (Pacific herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Proctacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8018, 30724;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=O.keta;
 RX MEDLINE=83195140; PubMed=6341999;
 RA Sherwood N., Elden L., Brownstein M., Spiess J., Rivier J., Vale W.;
 RT "Characterization of a teleost gonadotropin-releasing hormone."
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
 RN [2]
 RP SEQUENCE, AND FUNCTION.
 RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
 RA Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring."
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.

FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;
 Query Match 27.8%; Score 45; DB 1; Length 10;
 Best Local Similarity 77.8%; Pred. No. 0.57;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 III:II II
 DB 2 HWSYGLWPG 10
 III:II II
 RESULT 4
 GON2_CHICK STANDARD; PRT; 10 AA.
 AC P37043; P20408; P81750;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GNRH-II) (LH-RH II) (Luliberin II).
 OS Gallus gallus (Chicken).
 OS Alligator mississippiensis (American alligator),
 OS Squalus acanthias (Spiny dogfish),
 OS Hydrolagus colliei (Spotted ratfish) (Pacific ratfish), and
 OS Clupea pallasi (Pacific herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Chicken; TISSUE=Hypothalamus;
 RX MEDLINE=84222059; PubMed=6427779;
 RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
 RA Matsuo H.;
 RT "Identification of the second gonadotropin-releasing hormone in
 chicken hypothalamus: evidence that gonadotropin secretion is
 probably controlled by two distinct gonadotropin-releasing hormones
 in avian species."
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=A.mississippiensis; TISSUE=Brain;
 RX MEDLINE=91352338; PubMed=1882082;
 RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
 RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure of two forms of gonadotropin-releasing hormone
 from brains of the American alligator (Alligator mississippiensis)."
 RL Regul. Pept. 33:105-116(1991).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=S.acanthias; TISSUE=Brain;
 RX MEDLINE=92335300; PubMed=1631133;
 RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
 RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
 RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
 dogfish brain provides insight into GNRH evolution."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=H.colliei; TISSUE=Brain;
 RX MEDLINE=91340067; PubMed=1678723;
 RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
 RA Lee T.;
 RT "Primary structure of gonadotropin-releasing hormone from the brain
 of a holocephalan (ratfish: Hydrolagus colliei)."
 RL Gen. Comp. Endocrinol. 82:152-161(1991).
 RN [5]
 RP SEQUENCE, AND FUNCTION.
 RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

```

RA Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
RT hormones, including a novel form, from an ancient teleost, herring.";
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: B60066; RHAQ2.
DR PIR: A61126; A61126.
DR PIR: B46030; B46030.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;

Query Match 24.7%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWPFG 10

RESULT 5
GONL_SQUAC STANDARD; PRT; 10 AA.
AC P27429;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberein (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
DE (Luliberin).
OS Squalus acanthias (Spiny dogfish).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
CC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
TX TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Namvongchon S., Craig A.G.,
RA Nahrniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
RT dogfish brain provides insight into GNRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;

Query Match 24.7%; Score 40; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHGWPFG 10

RESULT 6
GON3_PETMA STANDARD; PRT; 10 AA.
AC P30948;

```

```

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberein III (Gonadotropin-releasing hormone III) (GNRH-III)
DE (Luliberin III).
OS Petromyzon marinus (Sea lamprey).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
CC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
TX TISSUE=Brain;
RX MEDLINE=93178316; PubMed=8440174;
RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
RT "Primary structure and biological activity of a third gonadotropin-
RT releasing hormone from lamprey brain.";
RL Endocrinology 132:1125-1131(1993).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1277 MW; 284B36237A1F5A3 CRC64;

Query Match 22.8%; Score 37; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 9.1;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSHDWKPFG 10

RESULT 7
GONL_CHEPR STANDARD; PRT; 10 AA.
AC P80677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberein I (Gonadotropin-releasing hormone I) (GNRH-I)
DE (Luliberin I).
OS Chelyosoma productum.
CC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
CC Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
RT and the evolutionary implications.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
CC THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -!- MASS SPECTROMETRY: MW=1246.56; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation.
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
SQ SEQUENCE 10 AA; 1264 MW; 284B3639DB5AB5A3 CRC64;

```

```
Query Match      19.8%; Score 32; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
   ||| :||
Db 2 HWSDFKPG 10

RESULT 8
GON2_CHEPR      STANDARD;      PRT;      10 AA.
AC P80678;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GNRH-II)
DE (luliberin II).
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
RT and the evolutionary implications.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
CC THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -!- MASS SPECTROMETRY: MW=1117.52; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 6 6 INTERCHAIN.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
FT MOD_RES 10 10 284B38DIEB735A3 CRC64;
SQ SEQUENCE 10 AA; 1135 MW; 284B38DIEB735A3 CRC64;

Query Match      17.3%; Score 28; DB 1; Length 10;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
   ||| ||
Db 2 HWSLCHAPG 10

RESULT 9
NRLA_ACISP      STANDARD;      PRT;      21 AA.
AC P33036;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Nitrilase (EC 3.5.5.1) (Fragment).
OS Acinetobacter sp. (strain AK226).
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=472;
RN [1]
RP SEQUENCE.
RX MEDLINE=91345837; PubMed=1369128;
RA Yamamoto K., Komatsu K.;

"Purification and characterization of nitrilase responsible for the
enantioselective hydrolysis from Acinetobacter sp. AK 226.";
RL Agric. Biol. Chem. 55:1459-1466(1991).
CC -!- FUNCTION: ACTS ON MANY KINDS OF NITRILE COMPOUNDS, SUCH AS
CC ALIPHATIC, AROMATIC, AND HETEROCYCLIC MONONITRILES OR DINITRILES.
CC PREFERS S-(-)-2-(4'-ISOBUTYLPHENYL)-PROPIONITRILE TO R-(+)-2-(4'-
CC ISOBUTYLPHENYL)-PROPIONITRILE AS THE SUBSTRATE.
CC -!- CATALYTIC ACTIVITY: A nitrile + H(2)O = a carboxylate + NH(3).
CC -!- SIMILARITY: BELONGS TO THE NITRILASE FAMILY.
DR PIR: J00386; J00386.
DR InterPro: IPR000132; Nitril_cyn_hydrolase.
DR PROSITE: PS00920; NITRIL_CHT_1; PARTIAL.
DR PROSITE: PS00921; NITRIL_CHT_2; PARTIAL.
KW Hydrolase.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2223 MW; 5FA741C41EAC619B CRC64;

Query Match      17.3%; Score 28; DB 1; Length 21;
Best Local Similarity 45.5%; Pred. No. 4.7e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 19 IKANSKFGIT 29
   : ||||| :
Db 1 VSYNSKFLAAT 11

RESULT 10
PSAF_SYNVU      STANDARD;      PRT;      27 AA.
AC P23078;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem I reaction centre subunit III (PSI-F) (Fragment).
GN PSF.
OS Synechococcus vulcanus (Thermosynechococcus vulcanus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32053;
RN [1]
RP SEQUENCE.
RX MEDLINE=89338747; PubMed=2503399;
RA Koike H., Ikeuchi M., Hiyyama T., Inoue Y.;
RT "Identification of photosystem I components from the cyanobacterium,
RT Synechococcus vulcanus by N-terminal sequencing.";
RL FEBS Lett. 253:257-263(1989).
CC -!- FUNCTION: PROBABLY PARTICIPATES IN EFFICIENCY OF ELECTRON TRANSFER
CC FROM PLASTOCYANIN TO P700 (OR CYTOCHROME C553 IN ALGAE AND
CC CYANOBACTERIA). THIS PLASTOCYANIN-DOCKING PROTEIN CONTRIBUTES
CC TO THE SPECIFIC ASSOCIATION OF PLASTOCYANIN TO PS I.
CC -!- SIMILARITY: BELONGS TO THE PSF FAMILY.
DR PIR: S05219; S05219.
KW Photosynthesis; Photosystem I.
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2685 MW; 30FFACDC3F258CA4 CRC64;

Query Match      16.7%; Score 27; DB 1; Length 27;
Best Local Similarity 41.7%; Pred. No. 8.9e+02;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 6 GLRPGSSGSLQ 17
   ||| :||
Db 4 GLVPAKDSFAFQ 15

RESULT 11
TRP3_LEUMA      STANDARD;      PRT;      19 AA.
ID TRP3_LEUMA
AC P81735;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tachykinin-related peptide 3 (LemTRP 3).
OS Leucophaea maderae (Madeira cockroach).
```


OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE.
RC TISSUE=Midgut; PubMed=8897641;
RX MEDLINE=97053012; PubMed=8897641;
RA Muren J.E., Naessel D.R.;
RT "Isolation of five tachykinin-related peptides from the midgut of
RT the cockroach Leucophaea maderae: existence of N-terminally extended
RT isoforms.";
RL Regul. Pept. 65:185-196(1996).
CC -!- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC -!- TISSUE SPECIFICITY: MIDGUT.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 19 19 AMIDATION.
SQ SEQUENCE 19 AA; 1930 MW; 99B5471A011625E5 CRC64;

Query Match 16.4%; Score 26.5; DB 1; Length 19;
Best Local Similarity 42.1%; Pred. No. 7.1e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 9 PGSSGPSLOYIRANSKFIF 27
III I I I I
DB 6 PGSK-----KAPSGFLG 17

RESULT 12
TXA3 ANESU STANDARD; PRT; 27 AA.
ID TXA3 ANESU STANDARD; PRT; 27 AA.
AC P01535;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neurotoxin III (Toxin ATX-III).
OS Anemonia sulcata (Snake-locks sea anemone).
OC Eukaryota; Metazoa; Chidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthaeae; Actinilidae; Anemonia.
OX NCBI_TaxID=6108;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=78084776; PubMed=23311;
RA Martinez G., Kopeyan C., Schweitz H., Lazdunski M.;
RT "Toxin III from Anemonia sulcata: primary structure.";
RL FEBS Lett. 84:247-252(1977).
RN [2]
RP PRELIMINARY SEQUENCE.
RX MEDLINE=78044787; PubMed=21843;
RA Beress L., Wunderer G., Wachter E.;
RT "Amino acid sequence of toxin III from Anemonia sulcata.";
RL Hoppe-Seyler's Z. Physiol. Chem. 358:985-988(1977).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=93343891; PubMed=8102051;
RA Norton R.S., Cross K., Braach-Maksvytis V., Wachter E.;
RT "1H-NMR study of the solution properties and secondary structure of
RT neurotoxin III from the sea anemone Anemonia sulcata.";
RL Biochem. J. 293:545-551(1993).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE=95244415; PubMed=7727358;
RA Manoleras N., Norton R.S.;
RT "Three-dimensional structure in solution of neurotoxin III from the
RT sea anemone Anemonia sulcata.";
RL Biochemistry 33:11051-11061(1994).
CC -!- FUNCTION: BINDS SPECIFICALLY TO THE SODIUM CHANNEL.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SEA ANEMONE SHORT TOXIN FAMILY.
DR PIR; A01798; TZA23.

DR PDB; IANS; 31-AUG-94.
KW Toxin; Neurotoxin; Sodium channel inhibitor; 3D-structure.
FT DISULFID 3 17
FT DISULFID 4 11
FT DISULFID 6 22
FT CONFLICT 22 23 CS -> SC (IN REF. 2).
SQ SEQUENCE 27 AA; 2938 MW; AA4E261FFAF34A7A CRC64;

Query Match 16.4%; Score 26.5; DB 1; Length 27;
Best Local Similarity 40.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 3 WSYGLRP-GSSGPSL 16
I I I I I
DB 13 WQNCVPEGCGPKV 27

RESULT 13
L52_ADE07 STANDARD; PRT; 27 AA.
ID L52_ADE07 STANDARD; PRT; 27 AA.
AC P05663;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Late L1 52 kDa protein (Fragment).
OS Human adenovirus type 7.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GOMEN;
RX MEDLINE=83183660; PubMed=6301944;
RA Engler J.A., Hoppe M.S., van Bree M.P.;
RT "The nucleotide sequence of the genes encoded in early region 2b of
RT human adenovirus type 7.";
RL Gene 21:145-159(1983).
CC -!- FUNCTION: INVOLVED IN VIRION ASSEMBLY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03000; CAA26776.1; -.
KW Late protein.
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3108 MW; E5A8288B117BB533 CRC64;

Query Match 16.0%; Score 26; DB 1; Length 27;
Best Local Similarity 45.5%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 LRPSSGPSLQ 17
: I I I I
DB 8 MRPQQAPSQ 18

RESULT 14
GUN_SCHCO STANDARD; PRT; 28 AA.
ID GUN_SCHCO STANDARD; PRT; 28 AA.
AC P81190;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Endoglucanase (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)
DE (Fragment).
OS Schizophyllum commune (Bracket fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Stereales; Schizophyllaceae; Schizophyllum.
OX NCBI_TaxID=5334;

```

RN SEQUENCE.
RP MEDLINE=97459758; PubMed=9315718;
RA Clarke A.J., Drummelsmith J., Yaguchi M.;
RT "Identification of the catalytic nucleophile in the cellulase from
RT Schizophyllum commune and assignment of the enzyme to Family 5,
RL subtype 5 of the glycosidases.";
FEBS Lett. 414:359-361(1997).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose. Attached to the membrane by a lipid anchor
CC -1- SUBCELLULAR LOCATION: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC (probable).
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
DR InterPro: IPR001547; Glyco_hydro_F5.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; PARTIAL.
KW Cellulose degradation; Hydrolase; Glycosidase; Zymogen; Membrane;
KW Lipoprotein.
FT ACT_SITE 20 20 NUCLEOPHILE.
FT NON_TER 28 28
SQ SEQUENCE 28 AA; 2937 MW; B3F1C0C99C9950BE CRC64;

Query Match 16.0%; Score 26; DB 1; Length 28;
Best Local Similarity 38.5%; Pred. No. 1.3e+03;
Matches 5; Conservative 6; Mismatches 0; Indels 2; Gaps 1;

QY 17 QYIKANSK--FIG 27
DB ::::|::: |::|
7 ENLKANNQRCFLG 19

RESULT 15
PETL_OENHO STANDARD; PRT; 31 AA.
AC QSWTK4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome B6-F complex subunit VI (Cytochrome b6f complex subunit
DE petL).
GN PETL.
OS Oenothera hookeri (Hooker's evening primrose).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=85636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, Johansen;
RX MEDLINE=20309318; PubMed=10852478;
RA Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Maier R.M.,
RA Chiu W.L., Sears B.;
RT "Complete nucleotide sequence of the Oenothera elata plastid
RT chromosome, representing plastome I of the five distinguishable
RT Euenothera plastomes.";
RL Mol. Gen. Genet. 263:581-585(2000).
CC -1- FUNCTION: THE CYTOCHROME B6-F COMPLEX FUNCTIONS IN THE LINEAR
CC CROSS-MEMBRANE TRANSPORT OF ELECTRONS BETWEEN PHOTOSYSTEM II AND
CC I, AS WELL AS IN CYCLIC ELECTRON FLOW AROUND PHOTOSYSTEM I. PETL
CC IS IMPORTANT FOR PHOTOAUTOTROPHIC GROWTH AS WELL AS FOR ELECTRON
CC TRANSFER EFFICIENCY AND STABILITY OF THE CYTOCHROME B6-F COMPLEX
CC (By similarity).
CC -1- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE-ASSOCIATED (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE PETL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL: AJ271079; CAB67175.1; -
KW Electron transport; Chloroplast; Respiratory chain; Transmembrane;
KW Thylakoid.
FT TRANSMEM 4 24 POTENTIAL.
SQ SEQUENCE 31 AA; 3415 MW; A015C65D2F325493 CRC64;

Query Match 16.0%; Score 26; DB 1; Length 31;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 25 FIGITEL 31
DB ::::|::: |::|
22 FIGLTKI 28

Search completed: October 10, 2002, 16:46:30
Job time : 9.5 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:44:12 ; Search time 19 Seconds
(without alignments)
282.255 Million cell updates/sec

Title: US-09-848-834A-15

Perfect score: 162

Sequence: 1 XHWXYGLRPGSSGSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 14138

Minimum DB seq length: 0

Maximum DB seq length: 31

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	23.5	20	11 Q9EQX8	Q9eqx8 mus musculus
2	36	22.2	26	1 Q50835	Q50835 methanococc
3	33	20.4	30	11 Q62536	Q62536 mus spretus
4	33	20.4	31	2 Q9R9C1	Q9r9c1 borrelia bu
5	32.5	20.1	20	11 Q9QVF6	Q9qvf6 rattus sp.
6	32	19.8	24	10 Q9S937	Q9s937 beta vulgar
7	32	19.8	31	5 Q9TWK5	Q9twk5 mytilus edu
8	30.5	18.8	30	2 Q9R5S3	Q9r5s3 leuconostoc
9	30	18.5	23	2 Q9R5R3	Q9r5r3 photobacter
10	30	18.5	24	2 Q9ZEQ3	Q9zeh3 neisseria m
11	30	18.5	24	4 Q96H37	Q96h37 homo sapien
12	30	18.5	26	4 Q9BSM8	Q9bsm8 homo sapien
13	30	18.5	28	2 Q93JY7	Q93jy7 mycobacteri
14	30	18.5	30	2 Q45966	Q45966 coxiella bu
15	29	17.9	15	11 Q9QUZ3	Q9quz3 rattus sp.
16	29	17.9	18	13 Q90791	Q90791 gallus gall

17	29	17.9	20	11 Q9QUZ4	Q9quz4 rattus sp.
18	29	17.9	24	2 Q9R3Z7	Q9r3z7 neisseria m
19	29	17.9	24	4 Q9NQY4	Q9ncy4 homo sapien
20	29	17.9	27	12 Q37180	Q37180 hepatitis c
21	29	17.9	30	4 Q9UBS6	Q9ubs6 homo sapien
22	29	17.9	31	16 Q50992	Q50992 borrelia bu
23	28.5	17.6	23	4 Q96155	Q96155 homo sapien
24	28.5	17.6	30	4 Q9UMI6	Q9umi6 homo sapien
25	28.5	17.6	30	16 Q989K8	Q989k8 rhizobium 1
26	28	17.3	16	4 Q9UC54	Q9uc54 homo sapien
27	28	17.3	19	5 Q9TWJ8	Q9twj8 mytilus edu
28	28	17.3	20	6 Q95MK6	Q95mk6 eulemur cor
29	28	17.3	20	6 Q95MK5	Q95mk5 varecia var
30	28	17.3	23	5 Q9TWJ9	Q9twj9 mytilus edu
31	28	17.3	27	7 Q31210	Q31210 mus musculu
32	28	17.3	27	8 Q94Q58	Q94q58 schistosoma
33	28	17.3	27	12 Q37183	Q37183 hepatitis c
34	28	17.3	27	12 Q37184	Q37184 hepatitis c
35	28	17.3	28	4 Q9NQB8	Q9ncb8 homo sapien
36	27	16.7	12	4 Q13865	Q13865 homo sapien
37	27	16.7	17	11 Q97758	P97758 mus musculu
38	27	16.7	20	11 Q9QVF9	Q9qvf9 rattus sp.
39	27	16.7	21	10 Q9S8K6	Q9s8k6 mesembryant
40	27	16.7	23	13 Q90759	Q90759 gallus gall
41	27	16.7	24	2 Q9R558	Q9r558 streptococc
42	27	16.7	24	4 Q14877	Q14877 homo sapien
43	27	16.7	25	4 Q9UQU9	Q9uqu9 homo sapien
44	27	16.7	25	4 Q9UQB1	Q9ub1 homo sapien
45	27	16.7	28	10 Q94IR9	Q94ir9 pinus radia

ALIGNMENTS

RESULT 1

Q9EQX8 PRELIMINARY; PRT; 20 AA.
ID Q9EQX8;
AC Q9EQX8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE GLUTAMATE DECARBOXYLASE 65 (FRAGMENT).
GN GAD65.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV; TISSUE=LIVER;
RA Makinae K., Kobayashi T., Kobayashi T., Sinkawa H., Sakagami H.,
RA Kondo H., fashiro F., Miyazaki J., Obata K., Tamura S., Yanagawa Y.;
RT "Structure of the Mouse Glutamate Decarboxylase 65 Gene and its
RT Promoter";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AB032757; BAB20415.1; --
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 1389 MW; 28F544B0E4C3768D CRC64;

Query Match 23.5%; Score 38; DB 11; Length 20;
Best Local Similarity 50.0%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 WSYGLRPGSSGP 14

11:1111:

Db 9 WSFGSEGGSDAP 20

RESULT 2

ID Q50835 PRELIMINARY; PRT; 26 AA.
ID Q50835;
AC Q50835;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE VANNIELLI ORF547-HISA INTERGENIC REGION (FRAGMENT).
OS Methanococcus vannielii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2187;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88124191; PubMed=2829115;
RA Brown J.W., Thomm M., Beckler G.S., Frey G., Stetter K.O., Reeve J.N.;
RT "An archaeobacterial RNA polymerase binding site and transcription
RL Nucleic Acids Res. 16:135-150(1988).
DR EMBL; X07391; CAA30299.1; -;
FT NON_TER 1
SQ SEQUENCE 26 AA; 3073 MW; 307BF49A6549F6A0 CRC64;

Query Match 22.2%; Score 36; DB 1; Length 26;
Best Local Similarity 60.0%; Pred. No. 2.3e+02;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 10 GSSGSPSLOYIKANSK 24
I I I I I I I I I
Db 7 GHSGLVLOYIKEVER 21

RESULT 3

ID Q62536 PRELIMINARY; PRT; 30 AA.
AC Q62536;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE GAMMA CRYSTALLIN B (3) (FRAGMENT).
GN CRYGB.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRET/EL;
RX MEDLINE=94319082; PubMed=8043949;
RA Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maizaki Y.,
RA Nadeau J.H.;
RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL Mamm. Genome 5:349-355(1994).
CC -1- FUNCTION: CRYSTALLINS ARE THE DOMINANT STRUCTURAL COMPONENTS OF
CC THE VERTEBRATE EYE LENS.
CC -1- DOMAIN: HAS A TWO-DOMAIN BETA-STRUCTURE, FOLDED INTO FOUR VERY
CC SIMILAR GREEK KEY MOTIFS.
CC -1- MISCELLANEOUS: THERE ARE SIX DIFFERENT GAMMA CRYSTALLINS
CC IDENTIFIED IN MOUSE LENS.
CC -1- SIMILARITY: TO OTHER GAMMA AND BETA CRYSTALLINS.
DR EMBL; U05704; AAB60467.1; -;
DR HSSP; P02526; IAMM.
DR MGD; MGI:88522; Crgbp.
KW Eye lens protein; Multigene family; Duplication.
FT NON_TER 1
FT DOMAIN <1 >30 MOTIF 4.
FT NON_TER 30
SQ SEQUENCE 30 AA; 3622 MW; D348F9752BD4AD2 CRC64;

Query Match 20.4%; Score 33; DB 11; Length 30;
Best Local Similarity 45.0%; Pred. No. 7.4e+02;
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 5 YGLRPGSGSPSLOYIKANSK 24
- - I I I I I I I I I
Db 6 YLLRPGMYRYLDWGAANK 25

RESULT 4

ID Q9R9C1 PRELIMINARY; PRT; 31 AA.
AC Q9R9C1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PLASMID CP32-4, POSSIBLE PARTITION PROTEINS (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid cp32-4.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B31;
RX MEDLINE=969361033; PubMed=9695920;
RA Stevenson B., Casjens S., Rosa P.;
RT "Evidence of past recombination events among the genes encoding the
RL Erp antigens of Borrelia burgdorferi.";
RL Microbiology 144:1869-1879(1998).
DR EMBL; AF022481; AAC35446.1; -;
KW Plasmid.
FT NON_TER 1
SQ SEQUENCE 31 AA; 3816 MW; 747301F137443A49 CRC64;

Query Match 20.4%; Score 33; DB 2; Length 31;
Best Local Similarity 35.0%; Pred. No. 7.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 5 YGLRPGSGSPSLOYIKANSK 24
I I I I I I I I I
Db 12 YGKYPEQGGLIKWIKNLK 31

RESULT 5

ID Q9QVF6 PRELIMINARY; PRT; 20 AA.
AC Q9QVF6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CARBONIC ANHYDRASE IV, CA IV.
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=92147688; PubMed=1737787;
RA Waheed A., Zhu X.L., Sly W.S.;
RT "Membrane-associated carbonic anhydrase from rat lung. Purification,
RT characterization, tissue distribution, and comparison with carbonic
RT anhydrase IVs of other mammals.";
RL J. Biol. Chem. 267:3308-3311(1992).
SQ SEQUENCE 20 AA; 2275 MW; 6CB2AC1CD44DEB65 CRC64;

Query Match 20.1%; Score 32.5; DB 11; Length 20;
Best Local Similarity 44.4%; Pred. No. 5.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 5; Gaps 2;

QY 2 HWSYGLR---PGS--SGP 14
I I I I I I I I I
Db 3 HWXYETQAKEPNSXXSGP 20

RESULT 6

ID Q9S937 PRELIMINARY; PRT; 24 AA.
AC Q9S937;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE H(+)-TRANSLOCATING (PYROPHOSPHATE-ENERGIZED) INORGANIC PYROPHOSPHATASE

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 16.2949 Seconds
(without alignments)
211.311 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSGLRPGSSGLQVIKANSKFIGITEL 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	52.5	109	22	AA1980
2	83	51.2	216	21	AA1981
3	82.5	50.9	158	19	AA1982
4	81	50.0	116	21	AA1983
5	81	50.0	116	21	AA1984
6	80.5	49.7	158	19	AA1985
7	79	48.8	118	21	AA1986
8	79	48.8	118	21	AA1987
9	78	48.1	136	22	AA1988
10	78	48.1	136	22	AA1989
11	78	48.1	182	21	AA1990

12	77	47.5	173	21	AA1991	DNA encoding osteo
13	76	46.9	19	21	AA1992	HLA class II bindi
14	76	46.9	37	15	AA1993	Universal immunost
15	76	46.9	37	15	AA1994	Universal immunost
16	76	46.9	47	15	AA1995	LHRH-containing im
17	76	46.9	124	21	AA1996	Modified human int
18	76	46.9	124	21	AA1997	Modified human int
19	76	46.9	160	22	AA1998	Growth differentia
20	76	46.9	573	8	AA1999	Portion of B fragm
21	76	46.9	1315	22	AA2000	Clostridium tetani
22	75	46.3	25	21	AA2001	PSMpep007 - P2 ins
23	75	46.3	693	21	AA2002	Mutant human PSM a
24	75	46.3	750	21	AA2003	Mutant human prost
25	75	46.3	750	21	AA2004	Mutant human prost
26	74	45.7	15	11	AA2005	Tetanus toxin epit
27	74	45.7	15	11	AA2006	Universal T-cell e
28	74	45.7	15	18	AA2007	Tetanus toxoid uni
29	74	45.7	15	18	AA2008	Tetanus toxoid uni
30	74	45.7	15	19	AA2009	Tetanus toxin frag
31	74	45.7	15	19	AA2010	Universal helper f
32	74	45.7	15	20	AA2011	T-helper epitope f
33	74	45.7	15	20	AA2012	T-cell epitope pep
34	74	45.7	15	20	AA2013	Tetanus toxoid epi
35	74	45.7	15	21	AA2014	Tetanus P2 epitope
36	74	45.7	15	21	AA2015	Tetanus toxoid T.C
37	74	45.7	15	21	AA2016	Foreign epitope P2
38	74	45.7	15	21	AA2017	Amino acid sequenc
39	74	45.7	15	21	AA2018	Clostridium tetani
40	74	45.7	15	21	AA2019	Tetanus toxoid pro
41	74	45.7	15	22	AA2020	Clostridium tetani
42	74	45.7	15	22	AA2021	Vaccine related MH
43	74	45.7	15	22	AA2022	Amino acid sequenc
44	74	45.7	15	22	AA2023	Wild-type rT830 (t
45	74	45.7	15	22	AA2024	Tetanus Toxoid uni

ALIGNMENTS

RESULT 1
AAB20147
ID AAB20147 standard; Protein; 109 AA.
XX
AC AAB20147;
XX
DT 30-APR-2001 (first entry)
XX
DE Growth differentiation factor 8 AutoVac construct GDF-8 P2-3.
XX
KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Clostridium tetani.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..82 /note= "identical to residues 267-348 of human
FT GDF-8"
FT Region 83..97 /note= "tetanus toxoid P2 epitope"
FT Region 98..109 /note= "identical to residues 364-375 of human
FT GDF-8"
FT Misc-difference 73 /note= "Cys-73 may be substituted by Ser to avoid
FT disulfide bond formation"
FT Misc-difference 90..91 /note= "optionally replaced by Glu-Gly"
XX
PN WO200105820-A2.

```

XX PD 25-JAN-2001.
XX PF 20-JUL-2000; 2000WO-DK00413.
XX PR 20-JUL-1999; 99DK-0001014.
XX PR 26-JUL-1999; 99US-0145275.
XX PA (MEBI-) M & E BIOTECH AS.
XX PI Halkier T, Mouritsen S, Klysner S;
XX WPI; 2001-112680/12.
XX PT Increasing the muscle mass of animals used in meat production by down
XX PT regulating growth differentiation factor 8 (GDF-8) activity in the
XX PT animal through induction of anti-GDF-8 antibody production.
XX PS Example 1; Page 99; 110pp; English.
XX CC The present sequence is that of AutoVac construct GDF-8 P2-3,
XX CC comprising the 109 C-terminal amino acid residues of human
XX CC growth differentiation factor 8 (GDF-8) in which residues 83-97 are
XX CC replaced by the promiscuous tetanus toxin T-cell epitope P2 (see
XX CC AAB20143). It is an object of the invention to produce a
XX CC recombinant therapeutic vaccine that is capable of effecting
XX CC down-regulation of GDF-8 in order to increase the muscle growth
XX CC rate of farm animals. The vaccines (see AAB20145-53) are capable
XX CC of breaking autotolerance against autologous GDF-8. They comprise
XX CC the C-terminal portion of human GDF-8 in which a portion of the
XX CC native sequence is replaced by a T-cell epitope such as P2, with
XX CC minimal disturbance of the authentic 3-dimensional structure of
XX CC the protein. Nucleic acids encoding the GDF-8 variants can be used
XX CC for genetic immunisation of the animals. Down-regulation of GDF-8
XX CC activity can increase muscle mass by up to at least 45% in cattle,
XX CC pigs and poultry used for meat production, reducing the need for
XX CC antibiotic feed-additives. Anti-GDF8 vaccines can be used to
XX CC treat human diseases such as cancer cachexia where muscle atrophy is
XX CC pronounced and for patients suffering from acute and chronic heart
XX CC failure.
XX CC Sequence 109 AA;
XX CC
XX CC Query Match 52.5%; Score 85; DB 22; Length 109;
XX CC Best Local Similarity 63.3%; Pred. No. 7.7e-05;
XX CC Matches 19; Conservative 2; Mismatches 1; Indels 8; Gaps 1;
XX CC
XX CC QY 10 GSSGP-----SLQYIKANSKFIGITEL 31
XX CC ||||| :|||||||
XX CC Db 68 GSAGPCCTTKMSPIQYIKANSKFIGITEL 97
XX CC
XX CC RESULT 2
XX CC AAY92665
XX CC ID AAY92665 standard; Peptide; 216 AA.
XX CC
XX CC XX AAY92665;
XX CC
XX CC XX 10-AUG-2000 (first entry)
XX CC
XX CC DE MUC-1 analogue containing foreign epitopes.
XX CC
XX CC Mucin repeat; MUC-1 analogue; vaccination; self-protein; cancer;
XX CC cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;
XX CC cell-associated peptide antigen; foreign epitope.
XX CC
XX CC OS Homo sapiens.
XX CC
XX CC Key, Location/Qualifiers
XX CC FH Peptide 61..75
XX CC FT /label= P2
XX CC FT Peptide 136..156
XX CC FT /label= P30

```

```

FT XX /note= "q"
PN XX WO200020027-A2.
PD XX 13-APR-2000.
XX PF 05-OCT-1999; 99WO-DK00525.
XX PR 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX PA (MEBI-) M & E BIOTECH AS.
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX DR WPI; 2000-349917/30.
XX PT Inducing immune responses to weakly immunogenic, tumor associated
XX PT peptide antigens for the treatment of breast and prostate cancer
XX PS Example 4; Page -; 220pp; English.
XX CC This is an immunogenized MUC-1 analogue containing foreign epitopes P2
XX CC and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate
XX CC specific membrane antigen (hPSM) can be used in the claimed method as an
XX CC autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX CC antibody binding regions and cysteine residues involved in disulfide
XX CC bonds are preserved in the immunogenized forms (see features table). 10
XX CC regions suitable for the insertion of foreign T helper epitopes were
XX CC identified. The method is used for inducing immune responses against
XX CC weakly immunogenic cell-associated peptide antigens (PA) such as those
XX CC associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)
XX CC and/or fibroblast growth factor 8b (FGF8b). The method comprises
XX CC effecting simultaneous presentation by antigen producing cells (APCs) of
XX CC the animals immune system of: (1) at least 1 CTL (cytotoxic
XX CC T-lymphocyte) group derived from the PA and/or at least 1 B-cell group
XX CC derived from the cell-associated PA; and (2) at least 1 first T helper
XX CC cell group which is foreign to the animal. Analogues of human PSM, human
XX CC Her2 and human/murine FGF8b comprising a substantial part of all known
XX CC and predicted CTL and B-cell epitopes of the respective PA and including
XX CC at least one foreign T helper epitope are also claimed. The method is
XX CC used to treat prostate, prostate/breast or breast cancer when the PA is
XX CC human PSM, FGF8b and Her2, respectively.
XX CC Note: This sequence does not appear in the specification. It was made
XX CC using the mucin repeat sequence (AAY92664), P2 and P30 (AAY92625-26),
XX CC which appear on pages 220, 213 and 214 respectively, of the
XX CC specification.
XX CC Sequence 216 AA;
XX CC
XX CC Query Match 51.2%; Score 83; DB 21; Length 216;
XX CC Best Local Similarity 57.6%; Pred. No. 0.00034;
XX CC Matches 19; Conservative 2; Mismatches 2; Indels 10; Gaps 1;
XX CC
XX CC QY 9 PGSSGP-----SLQYIKANSKFIGITEL 31
XX CC ||||| : |||||
XX CC Db 43 PGSTAPPAGVTSAPDTRQYIKANSKFIGITEL 75
XX CC
XX CC RESULT 3
XX CC AAW81331
XX CC ID AAW81331 standard; Protein; 158 AA.
XX CC
XX CC AC AAW81331;
XX CC
XX CC XX 21-APR-1999 (first entry)
XX CC
XX CC XX TNF2-7, a TNF-alpha analogue.
XX CC
XX CC Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
XX CC vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
XX CC cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;

```

```

KW asthmã-
XX Synthetic.
OS Homo sapiens.
XX
XX PN WO9846642-A1.
XX
XX PD 22-OCT-1998.
XX
XX PF 15-APR-1998; 98WO-DK00157.
XX
XX PR 24-APR-1997; 97US-0044187.
XX
XX PR 15-APR-1997; 97DK-0000418.
XX
XX PA (FERR ) FARM LAB FERRING AS.
XX
XX PI Dalum I, Elsnor H, Jensen MR, Mouritsen S;
XX
XX WPI: 1998-594561/50.
XX
XX DR N-PSDB; AAV68420.
XX
XX Modified human tumour necrosis factor-alpha - comprises
PT immunodominant T cell epitope, useful in vaccines to treat or
PT prevent TNF-associated diseases, e.g. cancer
XX
XX PS Claim 13; Page 73; 134pp; English.
XX
XX The present sequence represents a modified human tumour necrosis
CC factor-alpha (TNF-alpha) analogue. The analogues have no residual
CC TNF activity and are immunogenic in a large proportion of the human
CC population (by using promiscuous epitopes). The TNF-alpha analogue
CC is able to generate, in humans, neutralizing antibodies to wild-type
CC human TNF alpha, has at least one fragment of TNF substituted by a
CC peptide containing an immunodominant T-cell epitope, and at least one
CC TNF-alpha B-cell epitope. The substitution causes a significant change
CC in the amino acid sequence of any one of the strands in the front
CC beta-sheet, any of the connecting loops or any of the B', I or D strands
CC in the back beta-sheet. The TNF-alpha analogues are used as vaccines for
CC treatment or prevention of diseases associated with excessive release
CC or activity of TNF-alpha, e.g. rheumatoid arthritis, Crohn's disease,
CC ulcerative colitis, cancer of any sort, disseminated sclerosis, diabetes,
CC psoriasis, osteoporosis and asthma.
XX
XX SQ Sequence 158 AA;

Query Match 50.9%; Score 82.5; DB 19; Length 158;
Best Local Similarity 65.6%; Pred. NO. 0.00028;
Matches 21; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 3 WSVGLRPGSGGPS---LQYKANSKFIGITEL 31
DB 60 YSOVLPKGGGCPSTHVLQYKANSKFIGITEL 91

RESULT 4
AAB45502
ID AAB45502 standard; Protein; 116 AA.
XX
XX AAB45502;
XX
XX 26-FEB-2001 (first entry)
XX
XX Modified murine interleukin-5 SEQ ID NO: 14.
XX
XX Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
XX cancer; eosinophilia; vaccine; allergic rhinitis.
XX
XX Mus musculus.
XX
XX Clostridium tetani.
XX
XX WO200065058-A1.
XX
XX 02-NOV-2000.
XX
XX Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,

```

```

XX 19-APR-2000; 2000WO-DK00205.
XX
XX 23-APR-1999; 99DK-0000552.
XX
XX PR 06-MAY-1999; 99US-0132811.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Klysner S;
XX
XX WPI: 2000-672791/65.
XX
XX Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -
XX
XX Example 2; Page 129-130; 172pp; English.
XX
XX The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.
XX
XX SQ Sequence 116 AA;

Query Match 50.0%; Score 81; DB 21; Length 116;
Best Local Similarity 76.2%; Pred. NO. 0.00032;
Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 SSGPSLQYIKANSKFIGITEL 31
DB 24 TSNETWQYIKANSKFIGITEL 44

RESULT 5
AAB45526
ID AAB45526 standard; Protein; 116 AA.
XX
XX AAB45526;
XX
XX 26-FEB-2001 (first entry)
XX
XX Modified murine interleukin-5 SEQ ID NO: 52.
XX
XX Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
XX cancer; eosinophilia; vaccine; allergic rhinitis.
XX
XX Mus musculus.
XX
XX Clostridium tetani.
XX
XX WO200065058-A1.
XX
XX 02-NOV-2000.
XX
XX 19-APR-2000; 2000WO-DK00205.
XX
XX 23-APR-1999; 99DK-0000552.
XX
XX PR 06-MAY-1999; 99US-0132811.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Klysner S;
XX
XX WPI: 2000-672791/65.
XX
XX N-PSDB; AAC68879.
XX
XX Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,

```

PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -
XX
PS Disclosure; Page 159-160; 172pp; English.
XX
CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5.
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.
XX
SQ Sequence 116 AA;
Query Match 50.0%; Score 81; DB 21; Length 116;
Best Local Similarity 76.2%; Pred. No. 0.00032;
Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 11 SSGPSLQYIKANSKFIGITEL 31
DB 24 TSNETWQYIKANSKFIGITEL 44
RESULT 6
AAW81327
ID AAW81327 standard; Protein; 158 AA.
XX
AC AAW81327;
XX
DT 21-APR-1999 (first entry)
XX
DE TNF2-1, a TNF-alpha analogue.
XX
KW Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
KW vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
KW cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
KW asthma.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9846642-Al.
XX
PD 22-OCT-1998.
XX
PF 15-APR-1998; 98WO-DK00157.
XX
PR 24-APR-1997; 97US-0044187.
PR 15-APR-1997; 97DK-0000418.
XX
PA (FERR) FARM LAB FERRING AS.
XX
PI Dalum I, Elsnor H, Jensen MR, Mouritsen S;
XX
DR WPI; 1998-594561/50.
DR N-PSDB; AAV68416.
XX
XX Modified human tumour necrosis factor-alpha - comprises
PT immunodominant T cell epitope, useful in vaccines to treat or
PT prevent TNF-associated diseases, e.g. cancer
XX
PS Example 1; Page 65-66; 134pp; English.
XX
CC The present sequence represents a modified human tumour necrosis
CC factor-alpha (TNF-alpha) analogue. The analogues have no residual
CC TNF activity and are immunogenic in a large proportion of the human
CC population (by using promiscuous epitopes). The TNF-alpha analogue
CC is able to generate, in humans, neutralizing antibodies to wild-type
CC human TNF alpha, has at least one fragment of TNF substituted by a
CC peptide containing an immunodominant T-cell epitope, and at least one
CC TNF-alpha B-cell epitope. The substitution causes a significant change

CC in the amino acid sequence of any one of the strands in the front
CC beta-sheet, any of the connecting loops or any of the B', I or D strands
CC in the back beta-sheet. The TNF-alpha analogues are used as vaccines for
CC treatment or prevention of diseases associated with excessive release
CC or activity of TNF-alpha, e.g. rheumatoid arthritis, Crohn's disease,
CC ulcerative colitis, cancer of any sort, disseminated sclerosis, diabetes,
CC psoriasis, osteoporosis and asthma.
XX
SQ Sequence 158 AA;
Query Match 49.7%; Score 80.5; DB 19; Length 158;
Best Local Similarity 76.0%; Pred. No. 0.00054;
Matches 19; Conservative 1; Mismatches 4; Indels 1; Gaps 1;
QY 7 LRPGSSGSLQYIKANSKFIGITEL 31
DB 2 VRSSRTPS-QYIKANSKFIGITEL 25
RESULT 7
AAB45491
ID AAB45491 standard; Protein; 118 AA.
XX
AC AAB45491;
XX
DT 26-FEB-2001 (first entry)
XX
DE Modified human interleukin-5 SEQ ID NO: 3.
XX
KW Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS Homo sapiens.
OS Clostridium tetani.
XX
PN WO200065058-Al.
XX
PD 02-NOV-2000.
XX
PF 19-APR-2000; 2000WO-DK00205.
XX
PR 23-APR-1999; 99DK-0000552.
PR 06-MAY-1999; 99US-0132811.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Klysner S;
XX
DR WPI; 2000-672791/65.
XX
PT Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -
XX
PS Example 2; Page 120; 172pp; English.
XX
CC The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.
XX
SQ Sequence 118 AA;
Query Match 48.8%; Score 79; DB 21; Length 118;
Best Local Similarity 94.1%; Pred. No. 0.00065;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 15 SLQYIKANSKFIGITEL 31


```

Db      30 TLQYIKANSKFIGITEL 46

RESULT 8
AAB45518
ID      AAB45518 standard; Protein; 118 AA.
XX
AC      AAB45518;
XX
DT      26-FEB-2001 (first entry)
XX
DE      Modified human Interleukin-5 SEQ ID NO: 36.
XX
KW      Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW      cancer; eosinophilia; vaccine; allergic rhinitis.
XX
OS      Homo sapiens.
OS      Clostridium tetani.
XX
PN      WO2000065058-A1.
XX
PD      02-NOV-2000.
XX
PF      19-APR-2000; 2000WO-DK00205.
XX
PR      23-APR-1999; 99DK-0000552.
PR      06-MAY-1999; 99US-0132811.
XX
PA      (MEBI-) M & E BIOTECH AS.
XX
PI      Klysner S;
XX
DR      WPI; 2000-672791/65.
DR      N-PSDB; AAC68871.
XX
PT      Down-regulating interleukin 5 (IL-5) activity in humans by
PT      administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT      prophylaxis or amelioration of asthma or other chronic allergic
PT      conditions -
XX
PS      Example 2; Page 146; 172pp; English.
XX
CC      The present invention is concerned with methods of treating asthma,
CC      eosinophilia, allergic rhinitis and other allergic diseases. These
CC      involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC      proteins and their coding sequences to down-regulate IL-5 activity and
CC      thus reduce eosinophil numbers. The allergic diseases may be treated
CC      using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC      it is possible that they may be used in the treatment of cancer and
CC      helminthic infections.
XX
SQ      Sequence 118 AA;

Query Match      48.8%; Score 79; DB 21; Length 118;
Best Local Similarity 94.1%; Pred. No. 0.00065;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      15 SLQYIKANSKFIGITEL 31
       :|||||
Db      30 TLQYIKANSKFIGITEL 46

RESULT 9
AAB46190
ID      AAB46190 standard; peptide; 72 AA.
XX
AC      AAB46190;
XX
DT      04-APR-2001 (first entry)
XX
DE      Tetanus toxoid epitope fusion construct #10.
XX

KW      Amyloid deposit; APP; Abeta; brain; human; clearing response; nootropic;
KW      Fc receptor mediated phagocytosis; immunogenic response; neuroprotective;
KW      amyloid precursor protein; Alzheimer's disease.
XX
OS      Clostridium tetani.
XX
PN      WO200072880-A2.
XX
PD      07-DEC-2000.
XX
PF      26-MAY-2000; 2000WO-US14810.
XX
PR      28-MAY-1999; 99US-0322289.
XX
PA      (NEUR-) NEURALAB LTD.
XX
PI      Schenk DB, Bard F, Vasquez NJ, Yednock T;
XX
DR      WPI; 2001-032104/04.
XX
PT      Preventing or treating a disease associated with amyloid deposits,
PT      especially Alzheimer's disease, comprises administering amyloid
PT      specific antibody -
XX
PS      Disclosure; Page 32; 143pp; English.
XX
CC      This invention describes a novel method of preventing or treating a
CC      disease associated with amyloid deposits of amyloid precursor protein
CC      (APP) Abeta fragments in the brain of a patient, which comprises
CC      administering to the patient: (a) an antibody that binds to Abeta, the
CC      antibody binds to an amyloid deposit and induces a clearing response (Fc
CC      receptor mediated phagocytosis) against it (b) a polypeptide containing
CC      an N-terminal segment of at least residues 1-5 of Abeta; or (c) an agent
CC      that induces an immunogenic response against residues 1-3 to 7-11 of
CC      Abeta. The products of the invention have nootropic and neuroprotective
CC      activity. The method is also useful for monitoring a course of treatment
CC      being administered to a patient e.g. active and passive immunization. The
CC      methods are useful for prophylactic and therapeutic treatment of
CC      Alzheimer's disease.
XX
SQ      Sequence 72 AA;

Query Match      48.1%; Score 78; DB 22; Length 72;
Best Local Similarity 68.0%; Pred. No. 0.00051;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      7 LRPSSGSPSLQYIKANSKFIGITEL 31
       :||:|||||
Db      27 MEKASSVFNQYIKANSKFIGITEL 51

RESULT 10
AAB49089
ID      AAB49089 standard; Protein; 136 AA.
XX
AC      AAB49089;
XX
DT      27-MAR-2001 (first entry)
XX
DE      Amyloid beta tetanus toxoid/HA/CS fusion protein, SEQ ID NO:25.
XX
KW      Amyloid disease; amyloid fibril deposition; amyloid plaque;
KW      immunogenic; antibody; vaccine; Alzheimer's disease;
KW      type 2 diabetes; reactive system amyloidosis;
KW      systemic senile amyloidosis; familial amyloid
KW      transmissible spongiform encephalopathy; Creutzfeld-Jacob disease; Kuru;
KW      haemodialysis-associated beta-2-microglobulin deposition;
KW      amyloid beta peptide; universal T-cell epitope; neuroprotective.
XX
OS      Chimeric - Homo sapiens.
OS      Chimeric - Clostridium tetani.
OS      Chimeric - Influenza virus.
OS      Chimeric - Plasmodium falciparum.

```

XX WO200072876-A2.
 PN
 XX
 PD 07-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US15239.
 XX
 PR 01-JUN-1999; 99US-0137010.
 XX
 PA (NEUR-) NEURALAB LTD.
 XX
 PI Schenk DB;
 XX
 DR WPI; 2001-070921/08.
 XX
 PT Pharmaceutical composition comprising immunogen against amyloid
 PT component such as fibril peptide or protein, or antibody against
 PT amyloid component useful for treating amyloid diseases or amyloidoses -
 XX
 PS Disclosure; Page 46; 140pp; English.
 XX
 CC The invention relates to a novel pharmaceutical composition for
 CC preventing or treating a disease characterised by amyloid fibril
 CC deposits (amyloid plaques) in a patient. The pharmaceutical composition
 CC comprises an agent that will induce an immune response against an amyloid
 CC component, or an antibody or antibody fragment that binds to an amyloid
 CC component. The invention also relates to a method for determining
 CC the prognosis of a patient undergoing treatment for an amyloid disorder
 CC which involves measuring a patient serum amount of immunoreactivity
 CC against a selected amyloid component. A patient serum immunoreactivity
 CC of at least four times a base line serum immunoreactivity control level
 CC indicates a prognosis of improved status with respect to the disorder.
 CC The pharmaceutical compositions of the invention are useful for treating
 CC a wide variety of disorders characterised by amyloid fibril deposition in
 CC a patient. Such disorders include Alzheimer's disease characterised by
 CC amyloid beta peptide fibril deposits; type 2 diabetes characterised by
 CC islet amyloid protein peptide (IAPP, amylin) fibrils; reactive systemic
 CC amyloidosis associated with systemic inflammatory diseases (e.g.,
 CC rheumatoid arthritis, osteomyelitis, tuberculosis) characterised by AA
 CC fibrils derived from serum amyloid A protein (ApoSAA); systemic senile
 CC amyloidosis and familial amyloid cardiomyopathy characterised by ATTR
 CC fibrils derived from transthyretin (TTR); transmissible spongiform
 CC encephalopathies (e.g. Creutzfeldt-Jakob disease, Kuru) characterised by
 CC prion protein deposits; and beta-2-microglobulin deposits which form as
 CC a result of long term haemodialysis treatment. The present sequence
 CC represents an immunogenic fusion protein comprising an amyloid beta
 CC peptide fused to a universal T-cell epitope which may be used in a
 CC composition to treat or prevent Alzheimer's disease.
 XX
 SQ Sequence 136 AA;
 Query Match 48.1%; Score 78; DB 22; Length 136;
 Best Local Similarity 68.0%; Pred. No. 0.0011;
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 7 LRPSSGSPSQYIKANSKFIGITEL 31
 Db : || :|||||||
 27 MEKASSVFNQYIKANSKFIGITEL 51
 RESULT 11
 ID AAY84424
 XX AAY84424 standard; Protein; 182 AA.
 AC AAY84424;
 XX
 DT 25-JUL-2000 (first entry)
 XX
 DE An osteoprotegerin ligand/tetanus toxoid P30 epitope fusion.
 XX
 KW Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;
 KW tumour necrosis factor receptor; type II transmembrane protein;
 KW osteoclast differentiation; CSF-1; osteoclast activator;

KW immune response; osteoporosis; bone resorption; ss.
 XX Synthetic.
 OS Clostridium tetani.
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..14
 FT Protein 15..112
 FT /note= "His tag"
 FT /note= "residues 158-255 of murine OPGL"
 FT Peptide 113..127
 FT /note= "tetanus toxoid P2 epitope"
 FT Protein 128..182
 FT /note= "residues 262-316 of murine OPGL"
 XX
 PN WO200015807-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 13-SEP-1999; 99WO-DK00481.
 XX
 PR 15-SEP-1998; 98DK-0001164.
 PR 02-OCT-1998; 98US-0102896.
 XX
 PA (MEBI-) M & E BIOTECH AS.
 XX
 PI Halkier T, Haaning J;
 XX
 DR WPI; 2000-271444/23.
 DR N-PSDB; AAZ99971.
 XX
 PT In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used
 PT to treat, prevent and ameliorate osteoporosis -
 XX
 PS Example; Page 97-98; 110pp; English.
 XX
 CC The present sequence encodes a fusion protein of murine osteoprotegerin
 CC ligand (OPGL) and tetanus toxoid P2 epitope. Osteoprotegerin is a
 CC secreted member of the tumour necrosis factor receptor family, which
 CC blocks osteoclastogenesis in a dose dependent manner. The OPGL protein
 CC is synthesised as a type II transmembrane protein. The murine and human
 CC OPGL polypeptides are 87% homologous. OPGL is a potent osteoclast
 CC differentiation factor when combined with CSF-1. It is not capable of
 CC inducing osteoclast differentiation in the absence of CSF-1. OPGL is
 CC also an activator of mature osteoclasts. The specification describes a
 CC method for the in vivo down-regulation of OPGL activity in an animal.
 CC The method comprises using at least one OPGL polypeptide or subsequence,
 CC and/or at least one OPGL analogue to induce an immune response in the
 CC animal. The method and OPGL polypeptide are useful for treating,
 CC preventing and ameliorating osteoporosis or other diseases or conditions
 CC characterised by excessive bone resorption.
 XX
 SQ Sequence 182 AA;
 Query Match 48.1%; Score 78; DB 21; Length 182;
 Best Local Similarity 81.0%; Pred. No. 0.0015;
 Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 11 SSGPSLQYIKANSKFIGITEL 31
 Db || :|||||||
 107 SSHNLMOYIKANSKFIGITEL 127
 RESULT 12
 ID AAY84425
 XX AAY84425 standard; Protein; 173 AA.
 AC AAY84425;
 XX
 DT 25-JUL-2000 (first entry)
 XX
 DE DNA encoding osteoprotegerin ligand/tetanus toxoid P30 epitope fusion.

XX Osteoprotegerin ligand; OPGL; osteoprotegerin; osteoclastogenesis;
 KW tumour necrosis factor receptor; type II transmembrane protein;
 KW osteoclast differentiation; CSF-1; osteoclast activator;
 KW immune response; osteoporosis; bone resorption.
 XX Synthetic.
 OS Clostridium tetani.
 OS Mus musculus.
 XX Key Location/Qualifiers
 FT Peptide 1..14
 FT Protein 15..144 /note= "His tag"
 FT Peptide /note= "residues 158-287 of murine OPGL"
 FT Peptide 145..159
 FT Protein /note= "tetanus toxoid P2 epitope"
 FT Protein 160..173
 FT /note= "residues 303-316 of murine OPGL"
 XX WO200015807-A1.
 PN 23-MAR-2000.
 XX 13-SEP-1999; 99WO-DK00481.
 XX 15-SEP-1998; 98DK-0001164.
 PR 02-OCT-1998; 98US-0102896.
 XX (MEBI-) M & E BIOTECH AS.
 XX Halkier T, Haaning J;
 DR WPI; 2000-271444/23.
 DR N-PSDB; AA299972.
 XX In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used
 PT to treat, prevent and ameliorate osteoporosis -
 XX Example; Page 99-100; 110pp; English.
 XX The present sequence represents fusion protein of murine osteoprotegerin
 CC ligand (OPGL) and tetanus toxoid P2 epitope. Osteoprotegerin is a
 CC secreted member of the tumour necrosis factor receptor family, which
 CC blocks osteoclastogenesis in a dose dependent manner. The OPGL protein
 CC is synthesised as a type II transmembrane protein. The murine and human
 CC OPGL polypeptides are 87% homologous. OPGL is a potent osteoclast
 CC differentiation factor when combined with CSF-1. It is not capable of
 CC inducing osteoclast differentiation in the absence of CSF-1. OPGL is
 CC also an activator of mature osteoclasts. The specification describes a
 CC method for the in vivo down-regulation of OPGL activity in an animal.
 CC The method comprises using at least one OPGL polypeptide or subsequence,
 CC and/or at least one OPGL analogue to induce an immune response in the
 CC animal. The method and OPGL polypeptide are useful for treating,
 CC preventing and ameliorating osteoporosis or other diseases or conditions
 CC characterised by excessive bone resorption.
 XX Sequence 173 AA;
 SQ Query Match 47.5%; Score 77; DB 21; Length 173;
 Best Local Similarity 66.7%; Pred. No. 0.002;
 Matches 18; Conservative 1; Mismatches 4; Indels 4; Gaps 1;
 QY 5 YGLRPGSSGSPSLQYIKANSKFIGITEL 31
 Db 137 FKLRAEE---QYIKANSKFIGITEL 159
 RESULT 13
 AAY99055
 ID AAY99055 standard; Peptide; 19 AA.
 XX AAY99055;
 AC

XX 07-AUG-2000 (first entry)
 XX HLA class II binding antigen epitope peptide #244.
 XX Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
 KW immune response; chronic viral disease; cancer; autoimmune disease;
 KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
 KW allograft rejection; allergy; Lyme disease; hepatitis; prostate cancer;
 KW glomerulonephritis; food hypersensitivity; malaria.
 XX Clostridium tetani.
 OS WO9961916-A1.
 PN 02-DEC-1999.
 XX 28-MAY-1999; 99WO-US12066.
 XX 29-MAY-1998; 98US-0087192.
 XX (EPIM-) EPIMUNE INC.
 PA Sette A, Southwood S, Sidney J;
 PI WPI; 2000-097143/08.
 DR New compositions containing immunogenic peptide epitopes for various
 PT HLA class II DR molecules useful for inducing helper T cell response -
 XX Claim 1; Page 44; 60pp; English.
 XX The present invention relates to a new pharmaceutical composition
 CC comprising a unit dose form of a peptide, or analogue, comprising an
 CC epitope selected from those represented by peptides AAY98812-Y99339
 CC which are derived from various antigens for various human leucocyte
 CC antigen class DR molecules, representative of the world wide population.
 CC The peptide/analogue binds to an HLA class II molecule at an IC-50 of
 CC less than or equal to 1,000 nM. The pharmaceutical can be used to induce
 CC a helper T cell response. The pharmaceutical focuses the immune response
 CC towards selected determinants and could therefore be used in cases of
 CC chronic viral diseases and cancer. Examples of diseases that can be
 CC treated using the peptide containing pharmaceutical include autoimmune
 CC diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia
 CC gravis), allograft rejection, allergies, Lyme disease, hepatitis,
 CC post-streptococcal endocarditis or glomerulonephritis and food
 CC hypersensitivities. The peptide epitopes can be used to enhance immune
 CC responses against other immunogens administered with the peptides.
 CC Diseases which can be treated using immunogenic mixtures include
 CC prostate cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma,
 CC cervical carcinoma, lymphoma, and condyloma acuminatum. The peptides may
 CC also be used to make monoclonal antibodies useful as potential diagnostic
 CC or therapeutic agents. The peptides may also be useful as diagnostic
 CC reagents, for example, to determine the susceptibility of an individual
 CC to a treatment regimen. Also, the peptides may be used to predict which
 CC individuals will be at substantial risk of developing chronic infection.
 CC The selection of appropriate T and B cell epitopes should allow the
 CC development of epitope based vaccines particularly towards conserved
 CC epitopes of pathogens which are characterized by high sequence
 CC variability such as HIV, HCV and Malaria.
 XX Sequence 19 AA;
 SQ Query Match 46.9%; Score 76; DB 21; Length 19;
 Best Local Similarity 93.8%; Pred. No. 0.00022;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 16 LQYIKANSKFIGITEL 31
 Db 2 MQYIKANSKFIGITEL 17
 RESULT 14

AAR65389
ID AAR65389 standard; peptide; 37 AA.

XX
AC AAR65389;

XX DT 21-SEP-1995 (first entry)

XX XX Universal immunostimulator having GG spacers.

XX KW Helper T cell epitope; universal immune stimulator; invasive; haptens;

XX KW tetanus toxin.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Domain 3...19 /note= "tetanus toxin helper T cell epitope"

XX FT Domain 22..37

XX FT /note= "invasin domain"

XX XX

XX PN WO9425060-A.

XX PD 10-NOV-1994.

XX PF 28-APR-1994; 94WO-US04832.

XX PR 27-APR-1993; 93US-0057166.

XX PR 14-APR-1994; 94US-0229275.

XX XX

XX PA (LADD/) LADD A E.

XX PA (WANG/) WANG C Y.

XX PA (ZAMB/) ZAMB T.

XX XX

XX PI Ladd AE, Wang CY, Zamb T;

XX XX WPI; 1994-357910/44.

XX PT Immunogenic luteinising hormone releasing hormone peptide(s) -

XX PT that suppress LHRH activity in males and females

XX PS Disclosure; Page 95; 213pp; English.

XX CC Synthetic immunogenic peptides are provided in which a universal immune

XX CC stimulator is linked to a peptide or protein haptens containing B cell

XX CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes

XX CC potent immune responses to the coupled peptide or protein. The

XX CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)

XX CC which elicits an immune response to the coupled peptide in members of

XX CC a heterogeneous population expressing diverse HLA phenotypes, and (B)

XX CC an adjuvant peptide sequence from the invasin protein of Yersinia.

XX CC spacer amino acid sequences (e.g. Gly-Gly) can be provided between the

XX CC invasin and Th domains and between the immune stimulator and haptens

XX CC components. When the haptens is LHRH, then optionally the invasin domain

XX CC can be omitted from the immune stimulator component.

XX CC The present sequence is an example of a -GG-Th-GG-invasin immune

XX CC stimulator to which a haptens can be bonded.

XX XX Sequence 37 AA;

XX Query Match 46.9%; Score 76; DB 15; Length 37;

XX Best Local Similarity 84.2%; Pred. No. 0.00047;

XX Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLOYIKANSKFIGITEL 31

DB 1 GGGKQYIKANSKFIGITEL 19

RESULT 15

AAR65383

ID AAR65383 standard; peptide; 37 AA.

XX

AC AAR65383;

XX 21-SEP-1995 (first entry)

XX XX Universal immunostimulator having GG spacers.

XX KW Helper T cell epitope; universal immune stimulator; invasive; haptens;

XX KW tetanus toxin.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Domain 1..16 /note= "invasin domain"

XX FT Domain 19..35

XX FT /note= "tetanus toxin helper T cell epitope"

XX XX

XX PN WO9425060-A.

XX PD 10-NOV-1994.

XX PF 28-APR-1994; 94WO-US04832.

XX PR 27-APR-1993; 93US-0057166.

XX PR 14-APR-1994; 94US-0229275.

XX XX

XX PA (LADD/) LADD A E.

XX PA (WANG/) WANG C Y.

XX PA (ZAMB/) ZAMB T.

XX XX

XX PI Ladd AE, Wang CY, Zamb T;

XX XX WPI; 1994-357910/44.

XX PT Immunogenic luteinising hormone releasing hormone peptide(s) -

XX PT that suppress LHRH activity in males and females

XX PS Disclosure; Page 95; 213pp; English.

XX CC Synthetic immunogenic peptides are provided in which a universal immune

XX CC stimulator is linked to a peptide or protein haptens containing B cell

XX CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes

XX CC potent immune responses to the coupled peptide or protein. The

XX CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)

XX CC which elicits an immune response to the coupled peptide in members of

XX CC a heterogeneous population expressing diverse HLA phenotypes, and (B)

XX CC an adjuvant peptide sequence from the invasin protein of Yersinia.

XX CC spacer amino acid sequences (e.g. Gly-Gly) can be provided between the

XX CC invasin and Th domains and between the immune stimulator and haptens

XX CC components. When the haptens is LHRH, then optionally the invasin domain

XX CC can be omitted from the immune stimulator component.

XX CC The present sequence is an example of an invasin-GG-Th-GG- immune

XX CC stimulator to which a haptens can be bonded.

XX XX Sequence 37 AA;

XX Query Match 46.9%; Score 76; DB 15; Length 37;

XX Best Local Similarity 84.2%; Pred. No. 0.00047;

XX Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLOYIKANSKFIGITEL 31

DB 17 GGGKQYIKANSKFIGITEL 35

Search completed: October 10, 2002, 16:05:13

Job time : 16.2949 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 9.94231 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-17

Perfect score: 248

Sequence: 1 XHWSYGLRPGSGPSLKLLS.....HRLGVGSPSLHWSYGLRPX 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	117	47.2	42	1	US-08-446-692-20
2	117	47.2	42	2	US-08-488-351A-20
3	116	46.8	27	1	US-08-446-692-19
4	116	46.8	27	2	US-08-488-351A-19
5	116	46.8	27	3	US-08-100-414B-36
6	116	46.8	27	4	US-08-303-323-36
7	116	46.8	45	1	US-08-446-692-33
8	116	46.8	45	2	US-08-488-351A-33
9	113	45.6	27	3	US-08-100-414B-43
10	113	45.6	27	4	US-08-303-323-43
11	110	44.4	31	3	US-08-100-414B-55
12	110	44.4	31	4	US-08-303-323-55
13	108	43.5	27	3	US-08-100-414B-41
14	108	43.5	27	3	US-08-100-414B-47
15	108	43.5	27	4	US-08-303-323-41
16	108	43.5	27	4	US-08-303-323-47
17	108	43.5	45	3	US-08-100-414B-45
18	108	43.5	45	4	US-08-303-323-45
19	107	43.1	31	3	US-08-100-414B-59
20	107	43.1	31	4	US-08-303-323-59
21	106	42.7	28	3	US-08-100-414B-38
22	106	42.7	28	4	US-08-303-323-38
23	106	42.7	40	2	US-08-460-502-10
24	105	42.3	31	3	US-08-100-414B-53
25	105	42.3	31	4	US-08-303-323-53
26	105	42.3	47	3	US-08-100-414B-60
27	105	42.3	47	4	US-08-303-323-60

28 105 42.3 49 3 US-09-100-414B-57 Sequence 57, Appl
29 105 42.3 49 4 US-09-303-323-57 Sequence 57, Appl
30 103.5 41.7 25 1 US-08-446-692-17 Sequence 17, Appl
31 103.5 41.7 25 2 US-08-488-351A-17 Sequence 17, Appl
32 103 41.5 27 3 US-09-100-414B-50 Sequence 50, Appl
33 103 41.5 27 4 US-09-303-323-50 Sequence 50, Appl
34 103 41.5 35 3 US-09-100-414B-80 Sequence 80, Appl
35 103 41.5 35 4 US-09-303-323-80 Sequence 80, Appl
36 101 40.7 40 2 US-08-460-502-11 Sequence 11, Appl
37 100 40.3 28 1 US-08-446-692-38 Sequence 38, Appl
38 100 40.3 28 2 US-08-488-351A-38 Sequence 38, Appl
39 100 40.3 46 1 US-08-446-692-40 Sequence 40, Appl
40 100 40.3 46 2 US-08-488-351A-40 Sequence 40, Appl
41 100 40.3 47 3 US-09-100-414B-63 Sequence 63, Appl
42 100 40.3 47 4 US-09-303-323-63 Sequence 63, Appl
43 99 39.9 28 3 US-09-100-414B-39 Sequence 39, Appl
44 99 39.9 28 4 US-09-303-323-39 Sequence 39, Appl
45 96.5 38.9 49 1 US-08-387-156-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-446-692-20
; Sequence 20, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-20

Query Match 47.2%; Score 117; DB 1; Length 42;

Best Local Similarity 82.8%; Pred No. 4, 9e-09;

Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 18 LLSEIKGVIVHRLEGVGPSPSLHWSYGLRP 46

Db 15 VLSEIKGVIVHRLEGVGPSPSLHWSYGLRP 41

```
RESULT 2
US-08-488-351A-20
; Sequence 20, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-20
Query Match 47.2%; Score 117; DB 2; Length 42;
Best Local Similarity 82.8%; Pred. No. 4.9e-09;
Matches 24; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 18 LLSKIGVIVHRLEGVEGPGSLHWSYGLRP 46
:|||||
DB 15 VLSKIGVIVHRLEGVGGE--HWSYGLRP 41

RESULT 3
US-08-446-692-19
; Sequence 19, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
```

```
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-19
Query Match 46.8%; Score 116; DB 1; Length 27;
Best Local Similarity 85.7%; Pred. No. 3.9e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVEGPGSLHWSYGLRP 46
:|||||
DB 1 LSEIKGVIVHRLEGVGGE--HWSYGLRP 26

RESULT 4
US-08-488-351A-19
; Sequence 19, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
```



```
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-33

Query Match 46.8%; Score 116; DB 1; Length 45;
Best Local Similarity 85.7%; Pred. No. 7.3e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | |
Db 19 LSEIKGVIVHRLEGVGE--HWSYGLRP 44

RESULT 8
US-08-488-351A-33
; Sequence 33, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-33

Query Match 46.8%; Score 116; DB 1; Length 45;
Best Local Similarity 85.7%; Pred. No. 7.3e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | |
Db 19 LSEIKGVIVHRLEGVGE--HWSYGLRP 44

RESULT 9
US-09-100-414B-43
; Sequence 43, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-43

Query Match 45.6%; Score 113; DB 3; Length 27;
Best Local Similarity 82.1%; Pred. No. 9.8e-09;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | |
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-33

Query Match 46.8%; Score 116; DB 2; Length 45;
Best Local Similarity 85.7%; Pred. No. 7.3e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | |
Db 19 LSEIKGVIVHRLEGVGE--HWSYGLRP 44

RESULT 9
US-09-100-414B-43
; Sequence 43, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-43

Query Match 45.6%; Score 113; DB 3; Length 27;
Best Local Similarity 82.1%; Pred. No. 9.8e-09;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
| | | | | | | | | | | | | | | | | | | | |
```


US-09-303-323-55

Query Match 44.4%; Score 110; DB 4; Length 31;
Best Local Similarity 78.6%; Pred. No. 2.9e-08;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGVGSPSLHWSYGLRP 46
|||||:|||||:|||||
Db 3 LSEIKGVIVHKLGVFGGEHWSYGLRP 30

RESULT 13

US-09-100-414B-41
; Sequence 41, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-41

Query Match 43.5%; Score 108; DB 3; Length 27;
Best Local Similarity 71.4%; Pred. No. 4.6e-08;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLGVGSPSLHWSYGLRP 46
|||||:|||||:|||||
Db 1 LSEIKGVIVHKLGVGIGE--HWSYGLRP 26

RESULT 14

US-09-100-414B-47
; Sequence 47, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York

STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-47

Query Match 43.5%; Score 108; DB 3; Length 27;
Best Local Similarity 75.0%; Pred. No. 4.6e-08;
Matches 21; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLGVGSPSLHWSYGLRP 46
|||||:|||||:|||||
Db 1 MSEIKGVIVHKLGVGGE--HWSYGLRP 26

RESULT 15

US-09-303-323-41
; Sequence 41, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:

```

; LENC2H: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-41

```

```

Query Match          43.5%  Score 108; DB 4; Length 27;
Best Local Similarity 71.4%  Pred. No. 4.6e-08;
Matches 20; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

```

```

Qy 19 LSEIKGVIVHRLGEGVPSLHWSYGLRP 46
    :|||||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 1 ISEIKGVIVHRLGEGVPSLHWSYGLRP 26

```

```

Search completed: October 10, 2002, 16:14:05
Job time : 9.94231 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 11.6453 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHSYGLRGSSGCSLFNPF.....VSASHLEGLSHWSYGLRPX 50

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	114	41.6	1315	1 BTCLTN	tentoxylisin (EC 3
2	67	24.5	1297	2 S39791	neurotoxin - Clost
3	65	23.7	1296	1 BTCLAB	bontoxilysin (EC 3
4	64	23.4	1268	2 S33411	botulinum neurotoxin
5	64	23.4	1291	2 I40631	non-proteolytic bo
6	63	23.0	369	2 S48109	neurotoxin type F
7	63	23.0	1274	2 I40813	neurotoxin type F
8	63	23.0	1291	1 A48940	bontoxilysin (EC 3
9	62.5	22.8	1291	2 S46431	botulinum neurotoxin
10	62.5	22.8	1291	2 A49777	botulinum neurotoxin
11	62	22.6	1296	2 I40645	botulinum neurotoxin
12	61.5	22.4	91	2 JCR7393	medaka-type gonado
13	61.5	22.4	92	1 RHUG	gonadoliberin prec
14	61	22.3	366	2 S48110	neurotoxin type F
15	61	22.3	502	2 T36389	probable transmemb
16	59.5	21.7	1285	2 S70582	botulinum neurotoxin
17	59.5	21.7	1999	2 AB2018	hypothetical prote
18	59	21.5	67	2 I78541	gonadoliberin prec
19	59	21.5	92	1 RHUG	gonadoliberin prec
20	58	21.2	10	1 RHGG	gonadoliberin - pi
21	58	21.2	10	1 RHSG	gonadoliberin - sh
22	58	21.2	89	2 I51423	gonadoliberin prec
23	58	21.2	90	1 RHMSG	gonadoliberin prec
24	57.5	21.0	352	1 VVVP24	coat protein VP2 -
25	57.5	21.0	836	2 JE0248	ATP-binding caset
26	57	20.8	665	2 H83403	hypothetical prote
27	56.5	20.6	98	2 I50739	gonadotropin-relea
28	56.5	20.6	345	2 A58519	hypothetical 345 p
29	56.5	20.6	367	2 S48106	neurotoxin type E

30	56.5	20.6	444	2 T240777	hypothetical prote
31	56.5	20.6	551	2 AC3572	lysyl-tRNA synthet
32	56.5	20.6	658	2 S06744	staphylocoagulase
33	56.5	20.6	1251	2 JH0256	botulinum neurotox
34	56.5	20.6	1252	2 S21178	botulinum neurotox
35	56	20.4	92	2 I50644	gonadoliberin I pr
36	56	20.4	141	2 S48593	probable methionyl
37	56	20.4	812	2 T01618	hypothetical prote
38	55.5	20.3	549	1 A47468	cytochrome-c oxida
39	55.5	20.3	658	2 A89786	staphylocoagulase
40	55.5	20.3	715	2 A41511	staphylocoagulase
41	55	20.1	496	2 T38197	probable myb-like
42	54.5	19.9	464	1 MNVWC	nonstructural prot
43	54.5	19.9	3122	2 T17202	DNA-directed DNA p
44	54.5	19.9	5232	2 A45086	HC-toxin synthetas
45	54	19.7	10	1 RHAQ1	gonadoliberin I -

ALIGNMENTS

RESULT 1

BTCLTN

tentoxylisin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin

C:Species: Clostridium tetani

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999

C:Accession: A25689; A25757; A25194; B25194; A60759; S69364; S09364

R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, EMBO J. 5, 2495-2502, 1986

A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b

A:Reference number: A25689; MUID:87053814

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770

R:Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A:Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; MUID:87040747

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774

A:Experimental source: strain CN3911

R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O. J. Bacteriol. 165, 21-27, 1986

A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C 1

A:Reference number: A25194; MUID:86085672

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <FA2>

A:Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921

A:Accession: B25194

A:Molecule type: protein

A:Residues: 865-894 <FA3>

R:Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T. Infect. Immun. 57, 3588-3593, 1989

A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termin

A:Reference number: A60759; MUID:90035436

A:Accession: A60759

A:Molecule type: protein

A:Residues: 461-475 <MAT>

R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Nlemann, H.; Widmann, C.; Corradin, G. J. Immunol. 142, 394-402, 1989

A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: JS0098; MUID:89093918

A:Contents: annotation: epitope region

R:Schiavo, G.; Benfenati, F.; Foulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta, Nature 359, 832-835, 1992

A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo

A:Reference number: S27125; MUID:93063293

A:Contents: annotation

A:Gene: atx; bota
 C:Function:
 A:Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associ
 C:Superfamily: tetanus toxin
 C:Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane p
 F:2-444/Product: bontoxilysin A light chain #status experimental <LGH
 F:445-1296/Product: bontoxilysin A heavy chain #status experimental <HWY>
 F:223,227/Binding site: zinc (His) #status predicted
 F:224/Active site: Glu #status predicted

Query Match	23.7%	Score 65;	DB 1;	Length 1296;
Best Local Similarity	56.2%	Pred. No. 6.9;		
Matches 9;	Conservative	5;	Mismatches 2;	Indels 0;
	Gaps	0;		

```
F:231/Active site: Glu #status predicted
```

Query Match 23.4%; Score 64; DB 2; Length 1291;
Best Local Similarity 62.5%; Pred. No. 9.3;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
I : I : |||:||||:|
Db 921 SMFLDFSVSWIRIPK 936

RESULT 6

S48109
neurotoxin type F - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C>Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
C:Accession: S48109
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48109
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-369 <CAM>
A:Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C:Superfamily: tetanus toxin

Query Match 23.0%; Score 63; DB 2; Length 369;
Best Local Similarity 56.2%; Pred. No. 2.9;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
I : |:|||||:|
Db 295 SRYQNFSISFWIRPK 310

RESULT 7

I40813
neurotoxin type F - Clostridium botulinum
C:Species: Clostridium botulinum
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40813; S48108
R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson,
FEMS Microbiol. Lett. 96, 225-230, 1992
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A:Reference number: I40644
A:Accession: I40813
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1274 <RES>
A:Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48108
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 634-1002 <CAM>
A:Cross-references: EMBL:X70816; NID:g407789; PIDN:CAA50147.1; PID:g407789
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 23.0%; Score 63; DB 2; Length 1274;
Best Local Similarity 56.2%; Pred. No. 12;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
I : |:|||||:|
Db . 928 SRYQNFSISFWIRPK 943

RESULT 8

A48940
bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
N:Alternate names: botulinum neurotoxin type B (BoNT/B)
C:Species: Clostridium botulinum
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C:Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A:Title: Molecular cloning of the Clostridium botulinum structural gene encoding the type B neurotoxin
A:Reference number: A48940; MUID:92384550
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1291 <WH>
A:Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
A:Experimental source: type B, Danish
A:Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIP:112081); this publication
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific isoforms
A:Reference number: S48103; MUID:94013372
A:Accession: S48105
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 634-994 <CAM>
A:Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
A:Experimental source: proteolytic type B, strain NCTC 7273
R:Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A:Description: Partial amino acid sequence of botulinum neurotoxin type B and comparison with type A
A:Reference number: S21575
A:Accession: S21575
A:Molecule type: DNA
A:Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZ>
A:Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384
R:Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Grebenstein, O.; Wernars, K.
J. Biol. Chem. 267, 14721-14729, 1992
A:Title: Minimal essential domains specifying toxicity of the light chains of tetanus toxin
A:Reference number: A42871; MUID:92340509
A:Accession: A42871
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-313, 'S', 315-451 <KUR>
A:Experimental source: strain Okra
A:Note: sequence extracted from NCBI backbone (NCBIP:109365)
R:DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A:Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with type A
A:Reference number: S07155; MUID:89000987
A:Accession: S07155
A:Molecule type: protein
A:Residues: 2-29, 'M', 31-45 <DAS>
A:Accession: S08562
A:Molecule type: protein
A:Residues: 442-463, 'R', 465-467 <DAS>
R:Schmidt, J.J.; Sathiyamoorthy, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A:Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A:Reference number: S07128; MUID:85197963
A:Accession: S07128
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-16 <SCH1>
A:Accession: S08573
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-17 <SCH2>
A:Accession: S08574
A:Status: preliminary
A:Molecule type: protein
A:Residues: 442-459 <SCH3>

R:Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, Nature 359, 832-835, 1992
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of syntaxin
A:Reference number: S27125; MUID:93063293
A:Contents: annotation
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses
C:Genes: Bot/b
C:Function: Bot/b
C:Superfamily: tetanus toxin
C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F:2-441/Product: bontoxilysin B light chain #status experimental <LGHT>
F:442-1291/Product: bontoxilysin B heavy chain #status experimental <HVY>
F:230-234/Binding site: zinc (His) #status predicted
F:231/Active site: Glu #status predicted

Query Match 23.0%; Score 63; DB 1; Length 1291;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSVFLRVPK 30
 I : I : I : I : I : I : I :
Db 921 SVFLDFSVFWIRPK 936

RESULT 9

S46431
botulinum neurotoxin C1 - Clostridium botulinum phage 1C (strain C 468)
N:Alternate names: BoNT/C1 protein
C:Species: Clostridium botulinum phage 1C
A:Variety: Strain C 468
C>Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 16-Jul-1999
C:Accession: S46431; S49107
R:Hauser, D.; Eklund, M.W.; Boquet, P.; Popoff, M.R.
Mol. Gen. Genet. 243, 631-640, 1994
A:Title: Organization of the botulinum neurotoxin C1 gene and its associated non-toxicity factor
A:Reference number: S46426; MUID:94301293
A:Accession: S46431
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1291 <HAU>
A:Cross-references: EMBL:X72793; NID:g516171; PIDN:CAA51313.1; PID:g516175
A:Experimental source: strain C 468
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C:Superfamily: tetanus toxin

Query Match 22.8%; Score 62.5; DB 2; Length 1291;
Best Local Similarity 25.4%; Pred. No. 14;
Matches 17; Conservative 13; Mismatches 12; Indels 25; Gaps 3;

QY 5 YGLRPGSSGP-----SLFNNFTVSVFLRVPKVSASHLEGPSL----- 41
 : : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 907 FDFKLGSSGDRGKVIVTQNEINIVNYSFISFIRINK-WSNLPGYTIDSVKNN 965

QY 42 -HWSYGL 47
 I : I :
Db 966 SGWSIGI 972

RESULT 10

A49777
botulinum neurotoxin type C1 precursor - Clostridium botulinum phage (type C, strain C:Species: Clostridium botulinum phage
C>Date: 10-Mar-1994 #sequence_revision 07-Apr-1994 #text_change 23-Mar-2001
C:Accession: S11291; A35396; S22166; A49777
R:Hauser, D.; Eklund, M.W.; Kurazono, H.; Binz, T.; Niemann, H.; Gill, D.M.; Boquet, Nucleic Acids Res. 18, 4924, 1990
A:Title: Nucleotide sequence of Clostridium botulinum C1 neurotoxin.
A:Reference number: S11291; MUID:90370487
A:Accession: S11291
A:Status: preliminary
A:Molecule type: DNA


```
A:Residues: 1-84, 'P', 86-1291 <HAU>
A:Cross-references: EMBL:X53751; NID:g14905; PIDN:CAA37780.1; PID:g14906
R:Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Takeshi, K.
Biochem. Biophys. Res. Commun. 171, 1304-1311, 1990
A:Title: The complete nucleotide sequence of the gene coding for botulinum type C-1 toxin
A:Reference number: A35396; MUID:91024998
A:Accession: A35396
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-669, 'R', 671-1291 <TS1>
R:Tsuzuki, K.; Kimura, K.; Fujii, N.; Yokosawa, N.; Oguma, K.
submitted to the EMBL Data Library, December 1991
A:Description: Nucleotide sequence of the gene for one of the components of hemagglutinin
A:Reference number: S22163
A:Accession: S22166
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1291 <TS2>
A:Cross-references: EMBL:X62389; NID:g558175; PIDN:CAA44263.1; PID:g40390
R:Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Oguma, K.
Appl. Environ. Microbiol. 57, 1168-1172, 1991
A:Title: Cloning of the structural gene for Clostridium botulinum type C-1 toxin and who
A:Reference number: A49777; MUID:91282468
A:Accession: A49777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-607 <TS3>
A:Cross-references: GB:D90210
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 22.88; Score 62.5; DB 2; Length 1291;
Best Local Similarity 25.48; Pred. No. 14;
Matches 17; Conservative 13; Mismatches 12; Indels 25; Gaps 3;

Qy 5 YGLRPGSSGP-----SLFNFTVSWFLRVPKVSASHLEGPSSL----- 41
Db 907 FDFKLGSSGDRGVIVTQENIYVNSWYFSFIRINK-WVSNLPGVTIDSVKNN 965

Qy 42 -HWSYGL 47
Db 966 SGWSIGI 972

RESULT 11
I40645
botulinum neurotoxin type A - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40645
R:Willens, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
Res. Microbiol. 144, 547-556, 1993
A:Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A
A:Reference number: I40645; MUID:94143603
A:Accession: I40645
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1296 <RES>
A:Cross-references: EMBL:X73423; NID:g507070; PIDN:CAA51824.1; PID:g507071
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 22.68; Score 62; DB 2; Length 1296;
Best Local Similarity 50.08; Pred. No. 17;
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 15 SLFNFTVSWFLRVPK 30
Db 936 SMYENFSTFWIKIPK 951

RESULT 12
JC7393
```

```
medaka-type gonadotropin-releasing hormone precursor - Japanese medaka
C:Species: Oryzias latipes (Japanese medaka)
C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: JC7393
R:Okubo, K.; Amano, M.; Yoshiura, Y.; Suetake, H.; Aida, K.
Biochem. Biophys. Res. Commun. 276, 298-303, 2000
A:Title: A novel form of gonadotropin-releasing hormone in the medaka, Oryzias latipes
A:Reference number: JC7393
A:Contents: Brain
A:Accession: JC7393
A:Molecule type: mRNA
A:Residues: 1-91 <OKU>
A:Cross-references: DDBJ:AB041333
C:Comment: This protein plays the roles as a hypophysiotropic factor, and a physiolog
C:Genetics:
A:Gene: mdgnrh
C:Keywords: brain

Query Match 22.48; Score 61.5; DB 2; Length 91;
Best Local Similarity 32.78; Pred. No. 0.92;
Matches 17; Conservative 7; Mismatches 23; Indels 5; Gaps 1;

Qy 2 HWSYGLRPGSSGSLFNFTVSWFLRV-----PKVSASHLEGPSSLHWSYGLR 48
Db 23 HWSFGLSPGGKRELKYPFTLENQIRLLNSNTPCSDLHLESSLAKIYRIK 74

RESULT 13
RHRTG
gonadoliberin precursor - rat
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hor
N:Contains: gonadoliberin; prolactin release-inhibiting factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A40147; B26173; A48410
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic
A:Reference number: A40147; MUID:89384661
A:Accession: A40147
A:Molecule type: DNA
A:Residues: 1-92 <BON>
A:Cross-references: GB:M31670; NID:g204447; PIDN:AAA41264.1; PID:g204448
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gona
A:Reference number: A94090; MUID:86094338
A:Accession: B26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing horm
A:Reference number: A48410; MUID:93105480
A:Accession: A48410
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <MAI>
A:Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
A:Experimental source: thymus
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)
C:Genetics:
A:Introns: 47/3; 79/3
C:Function:
A:Description: stimulates pituitary secretion of lutropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activi
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progadoliberin #status predicted <PGN>
F:24-33/Product: gonadoliberin #status predicted <GLN>
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 6.51709 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHWSYGLRPGSSGLENNF.....VSASHLEGLPSLHWSYGLRPX 50

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	41.6	1314	1	TETX_CLOTE
2	67	24.5	1296	1	BXG_CLOBO
3	65	23.7	1295	1	BXA1_CLOBO
4	63	23.0	1274	1	BXF_CLOBO
5	63	23.0	1290	1	BXB_CLOBO
6	62.5	22.8	1290	1	BXC1_CLOBO
7	62	22.6	1051	1	VP2_AHSV6
8	62	22.6	1295	1	BXA2_CLOBO
9	61.5	22.4	92	1	GON1_RAT
10	60	21.9	92	1	GON1_TUPGB
11	59	21.5	67	1	GON1_MACMU
12	59	21.5	92	1	GON1_HUMAN
13	58.5	21.4	1250	1	BXE_CLOBO
14	58	21.2	61	1	GON1_SHEEP
15	58	21.2	63	1	GON1_MESAU
16	58	21.2	89	1	GON1_XENLA
17	58	21.2	90	1	GON1_MOUSE
18	58	21.2	91	1	GON1_PIG
19	57.5	21.0	352	1	COA2_SV40
20	57.5	21.0	842	1	ABC6_HUMAN
21	56.5	20.6	94	1	GON1_HAPBU
22	56.5	20.6	658	1	STC1_STAAU
23	56.5	20.6	1250	1	BXE_CLOBO
24	56	20.4	92	1	GON1_CHICK
25	56	20.4	141	1	AMPX_MYCCA
26	56	20.4	469	1	G51_HUMAN
27	55.5	20.3	549	1	FIXN_BRAJA
28	55.5	20.3	715	1	STC2_STAAU
29	54.5	19.9	464	1	VNSS_TSWV1
30	54.5	19.9	3122	1	DPOZ_MOUSE
31	54.5	19.9	5217	1	HTS1_COCCA
32	54	19.7	10	1	GON1_ALLMI
33	54	19.7	585	1	YH70_SYNY3

ALIGNMENTS

RESULT 1

ID	TETX_CLOTE	STANDARD;	PRT;	1314 AA.
AC	P04958;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).			
OS	Clostridium tetani.			
OG	Plasmid.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1513;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87053814; PubMed=3536478;			
RA	Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,			
RA	Weller U., Hudel M., Habermann E., Niemann H.;			
RT	"Tetanus toxin: Primary structure, expression in E. coli, and			
RT	homology with botulinum toxins.";			
RL	EMBO J. 5:2495-2502(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CN3911;			
RX	MEDLINE=87040747; PubMed=3774547;			
RA	Fairweather N.F., Lyness V.A.;			
RT	"The complete nucleotide sequence of tetanus toxin.";			
RL	Nucleic Acids Res. 14:7809-7812(1986).			
RN	[3]			
RP	SEQUENCE OF 742-1314 FROM N.A.			
RX	MEDLINE=86085672; PubMed=3510187;			
RA	Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;			
RT	"Cloning, nucleotide sequencing, and expression of tetanus toxin			
RT	fragment C in Escherichia coli.";			
RL	J. Bacteriol. 165:21-27(1986).			
RN	[4]			
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.			
RX	MEDLINE=90201034; PubMed=2108021;			
RA	Kriegelstein K., Henschen A., Weller U., Habermann E.;			
RT	"Arrangement of disulfide bridges and positions of sulfhydryl groups			
RT	in tetanus toxin.";			
RL	Eur. J. Biochem. 188:39-45(1990).			
RN	[5]			
RP	PARTIAL SEQUENCE.			
RX	MEDLINE=92037649; PubMed=1935979;			
RA	Kriegelstein K.G., Henschen A.H., Weller U., Habermann E.;			
RT	"Limited proteolysis of tetanus toxin. Relation to activity and			
RT	identification of cleavage sites.";			
RL	Eur. J. Biochem. 202:41-51(1991).			
RN	[6]			
RP	IDENTIFICATION AS ZINC-PROTEASE.			
RX	MEDLINE=93010948; PubMed=1396558;			
RA	Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,			
RA	Montecucco C.;			
RT	"Tetanus toxin is a zinc protein and its inhibition of			
RT	neurotransmitter release and protease activity depend on zinc.";			

P47651 mycoplasma
P23671 clostridium
Q91a10 dicentrarch
Q01811 impatiens n
Q05098 pseudomonas
P46081 clostridium
O73812 morone saxa
P25472 clostridium
P70074 pagrus majo
P51919 sparus aura
Q9x4f5 rhizobium m
P26003 tomato spot

34 53.5 19.5 654 1 PSTA_MYCGE
35 53.5 19.5 760 1 AMY_CLOAB
36 53 19.3 99 1 GON1_DICLA
37 53 19.3 449 1 VNSS_INSYN
38 53 19.3 746 1 PFEA_PSEAE
39 53 19.3 1196 1 BXCN_CLOBO
40 52.5 19.2 95 1 GON1_MORSA
41 52.5 19.2 584 1 GOND_CLOCE
42 52 19.0 95 1 GON1_PAGNA
43 52 19.0 95 1 GON1_SPAAU
44 52 19.0 453 1 HGD_RHIME
45 52 19.0 467 1 VNSS_TSWVL

EMBO J. 11:3577-3583(1992).
 [7]
 RL IDENTIFICATION OF SUBSTRATE.
 RP Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 [8]
 RP X-RAY CRYSTALLOGRAPHY (2.7-9.3 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792(1997).
 CC -I- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
 CC BOND OF SYNAPTOSOMAL-2.
 CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-|-PHE-77 BOND IN
 CC SYNAPTOSOMAL-2.
 CC -I- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 CC AND ARE NON-TOXIC AFTER SEPARATION.
 CC -I- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 CC GANGLIOSIDE RECEPTORS.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X04436; CAA328033.1; -;
 CC EMBL; M12739; AAA32828.1; -;
 CC EMBL; X06214; CAA29564.1; -;
 CC PIR; A25689; BTCLTN.
 CC PDB; 1AF9; 29-APR-98.
 CC PDB; 1A8D; 14-OCT-98.
 CC MEROPS; M27.001; -;
 CC InterPro; IPR000395; Bontoxilysin.
 CC InterPro; IPR000130; Zn_Mtpeptidse.
 CC Pfam; PF01742; Peptidase_M27; 1.
 CC PRINTS; PR00760; BONTOXILYSIN.
 CC ProDom; PD001963; Bontoxilysin; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure.
 FT INIT_MET 0
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 233 233 BY SIMILARITY.
 FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
 FT TRANSMEM 226 246 POTENTIAL.
 FT TRANSMEM 669 689 POTENTIAL.
 FT DISULFID 438 466 INTERCHAIN.
 FT DISULFID 1076 1092
 SQ SEQUENCE 1314 AA; 150550 MW; 134C3657133EF81D CRC64;
 Query Match 41.6%; Score 114; DB 1; Length 1314;
 Best Local Similarity 95.5%; Pred. No. 4.8e-06;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFVSVFWLRPKVSASHLE 37

Db 945 MFNFTVSVFWLRPKVSASHLE 966
 RESULT 2
 BXG_CLOBO STANDARD; PRT; 1296 AA.
 ID BXG_CLOBO
 AC Q60393;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BoNT/G)
 DE (Bontoxilysin G).
 DE BOTG.
 OS Clostridium botulinum.
 GC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=113 / 30;
 RX MEDLINE=94092745; PubMed=8268233;
 RA Campbell K., Collins M.D., East A.K.;
 RT "Nucleotide sequence of the gene coding for Clostridium botulinum
 (Clostridium argentinense) type G neurotoxin: genealogical comparison
 with other clostridial neurotoxins.";
 RL Biochim. Biophys. Acta 1216:487-491(1993).
 CC -I- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE.
 CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -I- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -I- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -I- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X74162; CAA52275.1; -;
 CC HSSP; P10845; 3BTA.
 CC MEROPS; M27.002; -;
 CC InterPro; IPR000395; Bontoxilysin.
 CC InterPro; IPR000130; Zn_Mtpeptidse.
 CC Pfam; PF01742; Peptidase_M27; 1.
 CC PRINTS; PR00760; BONTOXILYSIN.
 CC ProDom; PD001963; Bontoxilysin; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC Neurotoxin; Hydrolase; Metalloprotease; Zinc.
 FT INIT_MET 0
 FT CHAIN 1 441 BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
 FT CHAIN 442 1296 BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
 FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 435 449 INTERCHAIN (PROBABLE).
 SQ SEQUENCE 1296 AA; 149013 MW; DC8E47E15F665C31 CRC64;
 Query Match 24.5%; Score 67; DB 1; Length 1296;

ches

15
927

LT 3

DR MEROPS; M27.002; -;
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_Mtpeptidase.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR PROSITE; PS001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT CHAIN 1 436 BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
FT CHAIN 437 1274 BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
FT METAL 227 227 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 228 228 BY SIMILARITY.
FT METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 429 445 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1274 AA; 146709 MW; 5899756A7438B921 CRC64;

Query Match 23.0%; Score 63; DB 1; Length 1274;
Best Local Similarity 56.2%; Pred. No. 6.9;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVFWLRVVK 30
I : ||::||::||
Db 928 SRYQNSIFSWRIKP 943

RESULT 5
BXB_CLOBO STANDARD; PRT; 1290 AA.
AC P10844; P10843;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BONT/B)
DE (Bontoxilysin B).
GN BOTB.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92384550; PubMed=1514783;
RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA Minton N.P.;
RT "Molecular cloning of the Clostridium botulinum structural gene
RT encoding the type B neurotoxin and determination of its entire
RT nucleotide sequence.";
RL Appl. Environ. Microbiol. 58:2345-2354(1992).
RN [2]
RP SEQUENCE OF 35-245 FROM N.A.
RC STRAIN=NTC 7273;
RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 633-993 FROM N.A.
RC STRAIN=NTC 7273;
RX MEDLINE=9401372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-44 AND 441-466.
RC STRAIN=657;
RX MEDLINE=89000987; PubMed=3139097;
RA Dasgupta B.R., Datta A.;
RT "Botulinum neurotoxin type B (strain 657): partial sequence and
RT similarity with tetanus toxin.";
RL Biochimie 70:811-817(1988).
RN [5]
RP SEQUENCE OF 1-16 AND 441-458.
RC STRAIN=OKRA;
RX MEDLINE=85197963; PubMed=3888113;

Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;
"Partial amino acid sequences of botulinum neurotoxins types B and
E.";
Arch. Biochem. Biophys. 238:544-548(1985).
[6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93054694; PubMed=1429690;
RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
"Botulinum neurotoxins are zinc proteins.";
J. Biol. Chem. 267:23479-23483(1992).
[7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin.";
Nature 359:832-835(1992).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF
SYNAPTOSOMAL-2.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81186; AAA23211.1; -;
DR EMBL; Z11934; CAA77991.1; -;
DR EMBL; X70817; CAA50148.1; -;
DR PIR; S07128; S07128.
DR PIR; S07155; S07155.
DR PIR; S08562; S08562.
DR PIR; S08573; S08573.
DR PIR; S08574; S08574.
DR PIR; A48940; A48940.
DR HSP; P10845; 3BTA.
DR MEROPS; M27.002; -;
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_Mtpeptidase.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR PROSITE; PS001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0
FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
FT CHAIN 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 230 230 BY SIMILARITY.
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 436 445 INTERCHAIN (PROBABLE).
FT CONFLICT 29 29 T -> M (IN REF. 4).
FT CONFLICT 217 217 R -> G (IN REF. 2).
FT CONFLICT 224 224 A -> S (IN REF. 2).

RT and between the orbiviruses.";
RL Virus Res. 53:53-73(1998).
CC -I- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
CC WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
CC MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC -I- SIMILARITY: BELONGS TO THE REOVIRUSES VP2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF021235; AAC04994.1; -
DR InterPro: IPR001742; Orbi_VP2.
DR Pfam; PF00898; Orbi_VP2; 2.
KW Coat protein.
SQ SEQUENCE 1051 AA; 12326 MW; 2804DB9E389F4B5F CRC64;

Query Match 22.6%; Score 62; DB 1; Length 1051;
Best Local Similarity 40.7%; Pred. No. 7.5;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 11 SSGPSLFNNFTVSEFWLRVPKVSASHLE 37
Db 630 TEGTVFSKRFVSYYVRVEXITTKHLE 656

RESULT 8
BXA2_CLOBO STANDARD; PRT; 1295 AA.
AC Q45894; P77780;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kyoto-F;
RX MEDLINE=94143603; PubMed=8310180;
RA Willens A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium
RT botulinum type A associated with infant botulism: comparison with
RT other clostridial neurotoxins.";
RL Res. Microbiol. 144:547-556(1993).
RN [2]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=Kyoto-F;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC -I- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC binds with high affinity to peripheral neuronal presynaptic
CC membrane, is then internalized by receptor-mediated endocytosis.
CC The C-terminus of the heavy chain (H) is responsible for the
CC adherence of the toxin to the cell surface while the N-terminus
CC mediates transport of the light chain from the endocytic vesicle
CC to the cytosol. After translocation, the light chain (L)
CC hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking
CC neurotransmitter release. Inhibition of acetylcholine release

CC results in flaccid paralysis, with frequent heart or respiratory
CC failure (By similarity).
CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -I- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H) (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X73423; CAA51824.1; -
DR EMBL; X87974; CAA61234.1; -
DR HSPSP; P10845; 3BTA.
DR MEROPS; M27.002; -
DR InterPro: IPR000395; Bontoxilysin.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0
FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 223 223 BY SIMILARITY.
FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 429 453 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1234 1279 BY SIMILARITY.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675 POTENTIAL.
SQ SEQUENCE 1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;

Query Match 22.6%; Score 62; DB 1; Length 1295;
Best Local Similarity 50.0%; Pred. No. 9.3;
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLENNFTVSEFWLRVPK 30
Db 935 SMYENFSTFWIKPK 950

RESULT 9
GONI_RAT
ID GONI_RAT STANDARD; PRT; 92 AA.
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor
DE I].
GN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting

```

RESULT 10
GONI_TUPGB STANDARD; PRT; 92 AA.
ID GONI_TUPGB
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (uterinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
DE GNRH1 OR GNRH.
OS Tupala glis bellangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OC NCBI_TaxID=9396;
RN [1]
RS SEQUENCE FROM N.A.
RP TISSUE-Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGnRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GnRH gene expression in a
RT placental mammal."
RT Gen. Comp. Endocrinol. 104:7-19(1996).
RL CC
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS: IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U63326; AAB16837.1; -.
CC InterPro; IPR002012; GnRH.
CC InterPro; IPR004079; GonadoliberinI.
CC Pfam; PF00446; GnRH; 1.
CC PRINTS; PR01541; GONADOLIBRNI.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Placenta; Signal.
CC FT SIGNAL 1 23 BY SIMILARITY.
CC FT CHAIN 24 92 PRONADOLIBERIN I.
CC FT PEPTIDE 24 33 GONADOLIBERIN I.
CC FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
CC FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC ACTIVITY.
CC FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
CC SIMILARITY).
CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
CC SIMILARITY).
CC SQ SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
Query Match 21.9%; Score 60; DB 1; Length 92;
Best Local Similarity 52.4%; Pred. No. 1;
Matches 11; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

Oy 2 HWSYGLRPGS--SGPSLPNNF 20
Db 25 HWSYGLRPGGKRNAENLIDSF 45

RESULT 11
GONI_MACMU STANDARD; PRT; 67 AA.
ID GONI_MACMU
AC P55247;

```



```

FT MOD_RES 24 24 ACTIVITY.
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 16 16 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT CONFLICT 16 16 W -> S (IN REF. 3).
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 51.5%; Score 59; DB 1; Length 92;
Best Local Similarity 52.4%; Pred. No. 1.4;
Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLFNNF 20
    |||||
Db 25 HWSYGLRPGKGRDAENLDSF 45

RESULT 13
BXE_CLOBO STANDARD; PRT; 1250 AA.
AC Q00496;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
DE (Bontoxilysin E).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium
OC NCBI_TaxID=1491;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BEUGA;
RX MEDLINE=92181428; PubMed=1543481;
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulin neurotoxin E derived from Clostridium
RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT ATCC 43181 and ATCC 43755).";
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=92174922; PubMed=1541280;
RA Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
RT "The complete amino acid sequence of the Clostridium botulinum type-E
RT neurotoxin, derived by nucleotide-sequence analysis of the encoding
RT gene.";
RL Eur. J. Biochem. 204:657-667(1992).
[3]
RN SEQUENCE OF 1-251 FROM N.A.
RP SEQUENCE=90264400; PubMed=2160960;
RA Binz T., Kurazono H., Willie M., Frevort J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
RT with other clostridial neurotoxins.";
RL J. Biol. Chem. 265:9153-9158(1990).
[4]
RN SEQUENCE OF 1-13.
RP MEDLINE=85197963; PubMed=3888113;
RA Schmidt J.J., Sathiamoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequences of botulinum neurotoxins types B and
RT E.";
RL Arch. Biochem. Biophys. 238:544-548(1985).
[5]
RN SEQUENCE OF 419-426.
RX MEDLINE=90344918; PubMed=2116911;
RA Gimenez J.A., Dasgupta B.R.;
RT "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
RT reveals the site trypsin nicks and homology with tetanus
RT neurotoxin.";
RL Biochimie 72:213-217(1990).
[6]
RN IDENTIFICATION OF SUBSTRATE.
RP MEDLINE=94063091; PubMed=8243676;
RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA Benfenati F., Wilson M.C., Montecucco C.;
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct

```

```

RT COOH-terminal peptide bonds.";
RL FEBS Lett. 335:99-103(1993).
RN IDENTIFICATION OF SUBSTRATE.
RP MEDLINE=94124495; PubMed=8294407;
RX Binz T., Biagi J., Yamaseki S., Baumeister A., Link E., Suedhof T.C.,
RA Jahn R., Niemann H.;
RT "Proteolysis of SNAP-25 by types E and A botulin neurotoxins.";
RL J. Biol. Chem. 269:1617-1620(1994).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
CC 181 BOND IN SNAP-25.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: X62089; CAA43999.1; -
EMBL: X62683; CAA44558.1; -
DR PIR: A60027; A60027.
DR PIR: B35294; B35294.
DR PIR: JH0257; JH0257.
DR PIR: S08575; S08575.
DR PIR: S18111; S18111.
DR PIR: S21178; S21178.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptidse.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 212 212 BY SIMILARITY.
FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 411 425 INTERCHAIN (PROBABLE).
FT CONFLICT 176 176 R -> G (IN REF. 2).
FT CONFLICT 197 197 C -> S (IN REF. 2 AND 3).
FT CONFLICT 339 339 R -> A (IN REF. 2).
FT CONFLICT 772 772 I -> L (IN REF. 2).
FT CONFLICT 962 962 FE -> LQ (IN REF. 2).
FT CONFLICT 966 966 R -> A (IN REF. 2).
FT CONFLICT 1194 1194 N -> NN (IN REF. 2).
SQ SEQUENCE 1250 AA; 143712 MW; D9FCE26DDA041EB4 CRC64;

Query Match 21.4%; Score 58.5; DB 1; Length 1250;
Best Local Similarity 23.6%; Pred. No. 24;
Matches 13; Conservative 9; Mismatches 4; Indels 2; Gaps 2;

```

Oy 17 FNNFTVFWLRVP-----KVSASHLEGPSLHWSY 45
:||||:||||:|
Db 911 YKNFSISFWVRIPYNDKIVNNEYYTIINCRDNNNGWKVSLNHNE---IIWTF 962

RESULT 14

GONL_SHEEP STANDARD; PRT; 61 AA.
ID GONL_SHEEP
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I]
DE (Fragment).
DE GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN-WESTERN RANGE; TISSUE-Hypothalamus;
RA Rodriguez R.E., Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE OF 1-10.
RX MEDLINE=72094314; PubMed=4550508;
RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vale W., Guillemin R.;
RT "Primary structure of the ovine hypothalamic luteinizing hormone-
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-Edman degradation).";
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U02517; AAA03433.1; -.
DR PIR; A93780; RHSHG.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; Gnrh; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >61 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;

Query Match 21.2%; Score 58; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10
|||||

RESULT 15

GONL_MESAU STANDARD; PRT; 63 AA.
ID GONL_MESAU
AC O09163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I]
DE (Fragment).
DE GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U091938; AAB51302.1; -.
DR InterPro; IPR002012; Gnrh.
DR Pfam; PF00446; Gnrh; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >63 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY
FT SIMILARITY).
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY (BY SIMILARITY).
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
FT NON_TER 63 63
SQ SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;

Query Match 21.2%; Score 58; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYGLRPG 10
|||||

Db 2 HWSYGLRPG 10

Search completed: October 10, 2002, 16:06:37
Job time : 7.51709 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 19.4444 Seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHSVYGLRPGSSGPSLENNF.....VSASHLEGLSLHWSYGLRPX 50

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mbc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	41.6	451	2 Q9LA13	Q9LA13 clostridium
2	114	41.6	1310	2 Q93N27	Q93N27 clostridium
3	64	23.4	361	2 Q45846	Q45846 clostridium
4	64	23.4	361	2 Q45848	Q45848 clostridium
5	64	23.4	441	2 Q9X708	Q9X708 clostridium
6	64	23.4	1268	2 Q45851	Q45851 clostridium
7	64	23.4	1291	2 Q9ZAJ8	Q9ZAJ8 clostridium
8	64	23.4	1291	2 Q08077	Q08077 clostridium
9	64	23.4	1291	2 Q93G71	Q93G71 clostridium
10	64	23.4	1291	2 Q93K0	Q93K0 clostridium
11	64	23.4	2082	2 Q9ACP4	Q9ACP4 streptomyces
12	62.5	22.8	1291	2 Q93HT3	Q93HT3 clostridium
13	62	22.6	1296	2 Q45894	Q45894 clostridium
14	61.5	22.4	91	13 Q9DGC8	Q9DGC8 oryzias lat
15	61	22.3	228	12 Q9J2H3	Q9J2H3 macaca mula
16	61	22.3	502	2 Q9X8T8	Q9X8T8 streptomyces

17	61	22.3	1278	2	Q57236	Q57236 clostridium
18	60.5	22.1	363	10	Q9FF71	Q9FF71 arabidopsis
19	59.5	21.7	657	10	Q9XE83	Q9XE83 sorghum bic
20	59.5	21.7	1285	2	Q45967	Q45967 clostridium
21	59.5	21.7	1285	2	Q9LBR1	Q9LBR1 clostridium
22	58	21.2	90	13	Q90Y63	Q90Y63 rana catesb
23	58	21.2	91	13	Q9PRH0	Q9PRH0 anguilla ja
24	57.5	21.0	234	12	Q92837	Q92837 simian viru
25	57.5	21.0	234	12	Q9W9A5	Q9W9A5 simian viru
26	57.5	21.0	352	12	Q98VM1	Q98VM1 simian viru
27	57.5	21.0	352	12	Q9W9F7	Q9W9F7 simian viru
28	57.5	21.0	352	12	Q910V5	Q910V5 simian viru
29	57.5	21.0	766	4	Q96ME8	Q96ME8 homo sapien
30	57.5	21.0	836	11	Q70595	Q70595 rattus norv
31	57.5	21.0	842	11	Q70DC39	Q70DC39 mus musculu
32	57.5	21.0	896	4	Q9HAQ7	Q9HAQ7 homo sapien
33	57	20.8	430	2	Q9XAV1	Q9XAV1 pseudomonas
34	57	20.8	665	16	Q912G5	Q912G5 pseudomonas
35	57	20.8	866	5	Q9VF20	Q9VF20 drosophila
36	57	20.8	1280	2	Q9ZAJ5	Q9ZAJ5 clostridium
37	56.5	20.6	113	2	Q9KW76	Q9KW76 staphylococ
38	56.5	20.6	140	2	Q9K2Q0	Q9K2Q0 staphylococ
39	56.5	20.6	140	2	Q9K2T0	Q9K2T0 staphylococ
40	56.5	20.6	167	2	Q9KW83	Q9KW83 staphylococ
41	56.5	20.6	167	2	Q9KW72	Q9KW72 staphylococ
42	56.5	20.6	167	2	Q9KW71	Q9KW71 staphylococ
43	56.5	20.6	194	2	Q9KW84	Q9KW84 staphylococ
44	56.5	20.6	322	2	Q9F3C9	Q9F3C9 streptomyce
45	56.5	20.6	367	2	Q45862	Q45862 clostridium

ALIGNMENTS

RESULT 1

Q9LA13 PRELIMINARY; PRT; 451 AA.

AC Q9LA13: 01-OCT-2000 (TREMREL. 15, Created)

DT 01-OCT-2000 (TREMREL. 15, Last sequence update)

DT 01-DEC-2001 (TREMREL. 19, Last annotation update)

DE TETANUS TOXIN (FRAGMENT).

OS Clostridium tetani.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1513;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=20886;

RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;

RT "Fragment C of Tetanus Toxin.";

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF154828; AAF73267.1; -.

DR HSP; P04958; 1A8D.

DR InterPro; IPR001064; Crystallin.

DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.

FT NON_TER

SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;

Query Match 41.6%; Score 114; DB 2; Length 451;
Best Local Similarity 95.5%; Pred. No. 7.7e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37

Db 82 MFNNFTVSFWLRVPKVSASHLE 103

RESULT 2

Q93N27 PRELIMINARY; PRT; 1310 AA.

ID Q93N27

AC Q93N27: 01-DEC-2001 (TREMREL. 19, Created)

```
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TETANUS TOXIN (FRAGMENT).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RQ Shumin Z., Dianliang L.;
RT "Cloning and sequence analysis of tetanus toxin gene.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF389424; AAK72964.2; -.
FT NON_TER 1
FT NON_TER 1310 1310
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 41.6%; Score 114; DB 2; Length 1310;
Best Local Similarity 95.5%; Pred. No. 2; 6e-06;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
:|||||
Db 947 MFNNFTVSFWLRVPKVSASHLE 968

RESULT 3
Q45846 PRELIMINARY; PRT; 361 AA.
AC Q45846;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RQ STRAIN=TYPE B, NON-PROTEOLYTIC 2129B (SCOTT);
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL: X70814; CAA50145.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 361 361
SQ SEQUENCE 361 AA; 42175 MW; 533EA98735CD98E1 CRC64;

Query Match 23.4%; Score 64; DB 2; Length 361;
Best Local Similarity 62.5%; Pred. No. 2.3;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
I:| :|:|:|:|:|
Db 288 SMFLDFSVFWIRIPK 303

RESULT 4
Q45848 PRELIMINARY; PRT; 361 AA.
AC Q45848;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN BONT/B.
OS Clostridium botulinum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RQ STRAIN=TYPE B, NON-PROTEOLYTIC EKLUND 2B (COLWORTH 229);
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL: X70819; CAA50150.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 361 361
SQ SEQUENCE 361 AA; 42131 MW; A2E0FFFC81F9533D CRC64;

Query Match 23.4%; Score 64; DB 2; Length 361;
Best Local Similarity 62.5%; Pred. No. 2.3;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
I:| :|:|:|:|:|
Db 288 SMFLDFSVFWIRIPK 303

RESULT 5
Q9X708 PRELIMINARY; PRT; 441 AA.
AC Q9X708;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RQ MEDLINE=93343691; PubMed=10413679;
RX Lalli G., Herreros J., Osborne S.I., Montecucco C., Rossetto O.,
RA Schiavo G.;
RT "Functional characterisation of tetanus and botulinum neurotoxins
binding domains.";
RL J. Cell Sci. 112:2715-2724(1999).
DR EMBL: AJ242628; CAB43706.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 441 441
SQ SEQUENCE 441 AA; 52772 MW; 721D0B468E8C95A4 CRC64;

Query Match 23.4%; Score 64; DB 2; Length 441;
Best Local Similarity 62.5%; Pred. No. 3;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
I:| :|:|:|:|:|
Db 79 SMFLDFSVFWIRIPK 94

RESULT 6
Q45851 PRELIMINARY; PRT; 1268 AA.
AC Q45851;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NEUROTOXIN TYPE F.
GN BONT /F.
```


OS Clostridium baratii.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1561;
 RN [1]

RP SEQUENCE FROM N.A.
 RC MEDLINE=93252228; PubMed=8486245;
 RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
 RA Richardson P.T.;
 RT "Nucleotide sequence of the gene coding for Clostridium baratii type F
 RT neurotoxin: Comparison with other clostridial neurotoxins";
 RL FEMS Microbiol. Lett. 108:175-182(1993).
 DR EMBL: X68262; CAA48329.1; -;
 DR HSP: P10845; 3BTA.
 DR MEROPS; M27.002; -;
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_Mtpeptidse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1268 AA; 145513 MW; 96304009IAC15ED2 CRC64;

Query Match 23.4%; Score 64; DB 2; Length 1268;
 Best Local Similarity 62.5%; Pred. No. 9.9;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVVK 30
 I : ||:||||:|
 Db 920 SRVNFVSFWIRPK 935

RESULT 7

Q92AJ8
 ID Q92AJ8 PRELIMINARY; PRT; 1291 AA.
 AC Q92AJ8;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE BONT PROTEIN.
 GN BONT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 3281 (ATCC 43757);
 RX MEDLINE=98440323; PubMed=9767710;
 RA Santos-Buelga J., Collins M.D., East A.K.;
 RT "Characterization of the genes encoding the Botulinum neurotoxin
 RT complex in a strain of clostridium botulinum producing type B & F
 RT neurotoxins";
 RL Curr. Microbiol. 37:312-318(1998).
 DR EMBL: Y13630; CAA73968.1; -;
 DR HSP: P10845; 3BTA.
 DR MEROPS; M27.002; -;
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_Mtpeptidse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1291 AA; 150840 MW; E4D3B0E46AB2E735 CRC64;

Query Match 23.4%; Score 64; DB 2; Length 1291;
 Best Local Similarity 62.5%; Pred. No. 10;
 Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVVK 30
 I : ||:||||:|
 Db 921 SMELDFSVFWIRPK 936

RESULT 8

Q08077
 ID Q08077 PRELIMINARY; PRT; 1291 AA.
 AC Q08077;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE BOTULINUM NEUROTOXIN TYPE B (EC 3.4.24.-) (BONT/B).
 GN BONT/B.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=ERLUND 17B ATCC25765;
 RX MEDLINE=94122659; PubMed=7764370;
 RA Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
 RT "Nucleotide sequence of the gene coding for non-proteolytic
 RT clostridium botulinum type B neurotoxin: comparison with other
 RT clostridial neurotoxins";
 RL Curr. Microbiol. 28:101-110(1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES SYNAPTOSOMAL-2.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS
 CC NEUROTOXIN.
 CC -1- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE
 CC REGION.
 DR EMBL: X71343; CAA50482.1; -;
 DR HSP: P10845; 3BTA.
 DR MEROPS; M27.002; -;
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_Mtpeptidse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 SQ SEQUENCE 1291 AA; 150513 MW; 71BCAF23D69FAA CRC64;

Query Match 23.4%; Score 64; DB 2; Length 1291;
 Best Local Similarity 62.5%; Pred. No. 10;
 Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVVK 30
 I : ||:||||:|
 Db 921 SMELDFSVFWIRPK 936

RESULT 9

Q93G71
 ID Q93G71 PRELIMINARY; PRT; 1291 AA.
 AC Q93G71;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE NEUROTOXIN TYPE B.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=1436;
RA  Kirma N., Ferreira J.L., Baumstark B.R.;
RT  "Characterization of six type A strains of Clostridium botulinum that
RT  contain type B toxin gene sequences.";
RL  Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AP295926; AA097132.1; -.
SQ  SEQUENCE 1291 AA; 150824 MW;  D7CA07BAE2EB8CD2 CRC64;

Query Match      23.4%; Score 64; DB 2; Length 1291;
Best Local Similarity 62.5%; Pred. No. 10;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY  15 SLFNNFTVSFWLRVPK 30
Db  921 SMFLDFSVSFWIRIPK 936
      | | :|:|:|:|:|
      | | :|:|:|:|:|

RESULT 10
Q933K0 ID Q933K0 PRELIMINARY; PRT; 1291 AA.
AC Q933K0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TYPE B CRYPTIC NEUROTOXIN.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=593, AND 588;
RA Kirma N., Ferreira J.L., Baumstark B.R.;
RT "Characterization of six type A strains of Clostridium botulinum that
RT contain type B toxin gene sequences.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF300466; AAL11499.1; -.
DR EMBL: AF300465; AAL11498.1; -.
SQ SEQUENCE 1291 AA; 150843 MW;  7AC1737B0FAS151 CRC64;

Query Match      23.4%; Score 64; DB 2; Length 1291;
Best Local Similarity 62.5%; Pred. No. 10;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY  15 SLFNNFTVSFWLRVPK 30
Db  921 SMFLDFSVSFWIRIPK 936
      | | :|:|:|:|:|
      | | :|:|:|:|:|

RESULT 11
Q9ACP4 ID Q9ACP4 PRELIMINARY; PRT; 2082 AA.
AC Q9ACP4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE LARGE SECRETED PROTEIN.
GN SCPI.274.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Murphy L.D., Harris D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
```

```

RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=98241550; PubMed=95731173;
RA Redenbach M., Ikeda K., Yamasaki M., Kinashi H.;
RT "Cloning and physical mapping of the EcoRI fragments of the giant
RT linear plasmid SCPI.";
RL J. Bacteriol. 180:2796-2799(1998).
DR EMBL: AL590464; CAC36800.1; -.
SQ SEQUENCE 2082 AA; 224419 MW;  3B66780C312665DB CRC64;

Query Match      23.4%; Score 64; DB 2; Length 2082;
Best Local Similarity 44.2%; Pred. No. 18;
Matches 19; Conservative 5; Mismatches 9; Indels 10; Gaps 3;

QY  3 WSYGLRPGSSGSPSLFNNFTVSFWLRVPKVSASHLEGSLHSY 45
Db  270 WSAG---GSSGA-----FTWSYGLDTPEVAG--EGFGLNLAY 302
      || | ||||
      || | ||||

RESULT 12
Q93HT3 ID Q93HT3 PRELIMINARY; PRT; 1291 AA.
AC Q93HT3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NEUROTOXIN.
GN NT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-YOICHI;
RX MEDLINE=21534265; PubMed=11676492;
RA Sagane Y., Kouguichi H., Watanabe T., Sunagawa H., Inoue K.,
RA Fujinaga Y., Oguma K., Ohyama T.;
RT "Role of C-Terminal Region of HA-33 Component of Botulinum Toxin in
RT Hemagglutination.";
RL Biochem. Biophys. Res. Commun. 288:650-657(2001).
DR EMBL: AB061780; BAB71749.1; -.
SQ SEQUENCE 1291 AA; 148869 MW;  4A21DB35B8743CF8 CRC64;

Query Match      22.8%; Score 62.5; DB 2; Length 1291;
Best Local Similarity 25.4%; Pred. No. 16;
Matches 17; Conservative 13; Mismatches 12; Indels 25; Gaps 3;

QY  5 YGLRPGSSGP-----SLFNNFTVSFWLRVPKVSASHLEGPSL----- 41
Db  907 FDFKLGSSEGRGKVIVTQENIVNYSFISFWIRINK-WVSNLPGYTIIDSVKN 965
      : : ||||
      : : ||||

RESULT 13
Q45894 ID Q45894 PRELIMINARY; PRT; 1296 AA.
AC Q45894; P77780;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE A (TYPE A NEUROTOXIN).
GN BONT OR AFX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 14.3889 Seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSVGLRPGSGPSLFNNFTVSFWLRPKVSASHLE 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL19:**
1: sp_archaea:**
2: sp_bacteria:**
3: sp_fungi:**
4: sp_human:**
5: sp_invertebrate:**
6: sp_mammal:**
7: sp_mhc:**
8: sp_organelle:**
9: sp_phage:**
10: sp_plant:**
11: sp_rodent:**
12: sp_virus:**
13: sp_vertebrate:**
14: sp_unclassified:**
15: sp_rvirus:**
16: sp_bacteriap:**
17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	57.0	451	2 Q9LA13	Q9LA13 clostridium
2	114	57.0	1310	2 Q93N27	Q93N27 clostridium
3	64	32.0	361	2 Q45846	Q45846 clostridium
4	64	32.0	361	2 Q45848	Q45848 clostridium
5	64	32.0	441	2 Q9X708	Q9X708 clostridium
6	64	32.0	1268	2 Q45851	Q45851 clostridium
7	64	32.0	1291	2 Q9ZAJ8	Q9ZAJ8 clostridium
8	64	32.0	1291	2 Q08077	Q08077 clostridium
9	64	32.0	1291	2 Q93G71	Q93G71 clostridium
10	64	32.0	1291	2 Q933K0	Q933K0 clostridium
11	62	31.0	1296	2 Q45894	Q45894 clostridium
12	61	30.5	1278	2 Q57236	Q57236 clostridium
13	58	29.0	90	13 Q90Y63	Q90Y63 rana catesb
14	58	29.0	91	13 Q9PRH0	Q9PRH0 anquilla ja
15	57.5	28.7	234	12 Q92837	Q92837 simian viru
16	57.5	28.7	234	12 Q9W9A5	Q9W9A5 simian viru

17	57.5	28.7	352	12 Q98VM1	Q98vm1 simian viru
18	57.5	28.7	352	12 Q9W9F7	Q9w9f7 simian viru
19	57.5	28.7	352	12 Q910V5	Q910v5 simian viru
20	57	28.5	1280	2 Q9ZAJ5	Q9Zaj5 clostridium
21	56	28.0	367	2 Q45862	Q45862 clostridium
22	56	28.0	367	2 Q45861	Q45861 clostridium
23	56	28.0	1251	2 Q9K395	Q9k395 clostridium
24	56	28.0	1255	2 Q9FAR6	Q9far6 clostridium
25	56	28.0	1291	2 Q93HT3	Q93ht3 clostridium
26	55	27.5	476	5 Q9N3R6	Q9n3r6 caenorhabdi
27	54.5	27.3	91	13 Q9DGC8	Q9dgc8 oryzias lat
28	53.5	26.8	251	16 Q92AM6	Q92am6 listeria in
29	53	26.5	1196	2 Q9LBS8	Q9lbs8 clostridium
30	53	26.5	1196	2 Q9LBR2	Q9lbr2 clostridium
31	53	26.5	1196	2 Q53550	Q53550 clostridium
32	53	26.5	1196	2 Q45916	Q45916 clostridium
33	53	26.5	1196	2 Q93HT4	Q93ht4 clostridium
34	53	26.5	1196	9 Q92X77	Q92x77 clostridium
35	53	26.5	1196	9 Q38197	Q38197 clostridium
36	53	26.5	1285	2 Q45967	Q45967 clostridium
37	53	26.5	1285	2 Q9LBR1	Q9lbr1 clostridium
38	52.5	26.2	609	12 Q98603	Q98603 peste-des-p
39	52	26.0	87	13 Q9YI26	Q9yi26 sparus aura
40	52	26.0	187	2 Q9EW11	Q9ew11 streptomyce
41	52	26.0	210	5 Q17476	Q17476 caenorhabdi
42	52	26.0	228	12 Q9J2H3	Q9j2h3 macaca mula
43	52	26.0	467	12 Q37367	Q37367 tomato spot
44	52	26.0	467	12 Q37369	Q37369 tomato spot
45	52	26.0	866	5 Q9VF20	Q9vf20 drosophila

ALIGNMENTS

RESULT 1
Q9LA13 ID Q9LA13 PRELIMINARY; PRT; 451 AA.
AC Q9LA13; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TETANUS TOXIN (FRAGMENT).
OS Clostridium tetani
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20886;
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
RT "Fragment C of Tetanus Toxin."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154828; AAF73267.1; -
DR HSSP; P04958; 1A8D.
DR InterPro; IPR001064; Crystallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;
Query Match 57.0%; Score 114; DB 2; Length 451;
Best Local Similarity 95.5%; Pred. No. 2.6e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 16 LFNNFTVSFWLRPKVSASHLE 37
:|||||
Db 82 MFNNFTVSFWLRPKVSASHLE 103
RESULT 2
Q93N27 ID Q93N27 PRELIMINARY; PRT; 1310 AA.
AC Q93N27
DT 01-DEC-2001 (TREMBLrel. 19, Created)

OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC	Clostridium.
OX	NCBI_TaxID=1491;
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=TYPE B, NON-PROTEOLYTIC EKLUND 2B (COLWORTH 229);
RC	MEDLINE=94013372; PubMed=8408542;
RX	Campbell K., East A.K., Collins M.D.;
FT	"Gene probes for identification of the botulinum neurotoxin gene and
RT	specific identification of neurotoxin types B, E, and F.";
RL	J. Clin. Microbiol. 31:2255-2262(1993).
DR	EMBL; X70819; CAA50150.1; -.
DR	HSSP; P10845; 3BTA.
KW	Neurotoxin.
FT	NON_TER 1 1
FT	NON_TER 361 361
SEQ	SEQUENCE 361 AA; 42131 MW; A2E0FFFC81F9533D CRC64;
Query Match	32.0%; Score 64; DB 2; Length 361;
Best Local Similarity	62.5%; Pred. No. 0.39;
Matches 10; Conservative	5; Mismatches 1; Indels 0; Gaps 0;
Qy	15 SLFNFTVSFWLRVPK 30
Db	1:1:1:1:1:1:1:1
288 SMFLDFSVFWIRPK 303	
RESULT 5	
O9X708	PRELIMINARY; PRT; 441 AA.
AC	O9X708;
DT	01-NOV-1999 (TRENBLrel. 12, Created)
DT	01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT	01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE	BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN	BONT/B.
OS	Clostridium botulinum.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC	Clostridium.
OX	NCBI_TaxID=1491;
[1]	
RN	SEQUENCE FROM N.A.
RP	MEDLINE=99343691; PubMed=10413679;
RX	Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
RA	Schiavo G.;
RT	"Functional characterisation of tetanus and botulinum neurotoxins
RL	binding domains.";
RL	J. Cell Sci. 112:2715-2724(1999).
DR	EMBL; AJ242628; CAB43706.1; -.
DR	HSSP; P10845; 3BTA.
KW	Neurotoxin.
FT	NON_TER 1 1
FT	NON_TER 441 441
SEQ	SEQUENCE 441 AA; 52772 MW; 721D0B468E8C95A4 CRC64;
Query Match	32.0%; Score 64; DB 2; Length 441;
Best Local Similarity	62.5%; Pred. No. 0.48;
Matches 10; Conservative	5; Mismatches 1; Indels 0; Gaps 0;
Qy	15 SLFNFTVSFWLRVPK 30
Db	1:1:1:1:1:1:1:1
79 SMFLDFSVFWIRPK 94	
RESULT 6	
Q45851	PRELIMINARY; PRT; 1268 AA.
ID	Q45851;
AC	Q45851;
DT	01-NOV-1996 (TRENBLrel. 01, Created)
DT	01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT	01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE	NEUROTOXIN TYPE F.
GN	BONT /F.

```

RESULT 8
Q08077
ID AC Q08077 PRELIMINARY; PRT; 1291 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (EC 3.4.24.-) (BONT/B).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=EKLUND 17B ATCC25765;
RX MEDLINE=941122659; PubMed=7764370;
RA Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT "Nucleotide sequence of the gene coding for non-proteolytic
RT clostridium botulinum type B neurotoxin: comparison with other
RT clostridial neurotoxins."
RL Curr. Microbiol. 28:101-110(1994).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTANSITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTANSITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CLEAVES SYNAPTOBREVIN-2.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS
CC NEUROTOXIN.
CC -1- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE
CC REGION.
CC EMBL; X71343; CAA50482.1; -.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
SQ SEQUENCE 1291 AA; 150513 MW; 71BCAFE23D69FAAA CRC64;

Query Match 32.08; Score 64; DB 2; Length 1291;
Best Local Similarity 62.58; Pred. No. 1.6;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 15 SLFNNFTVSFWLRVPK 30
|:|:|:|:|:|:|:|
Db 921 SMFLDFSVFWIRPK 936

RESULT 9
Q93G71
ID AC Q93G71 PRELIMINARY; PRT; 1291 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEUROTOXIN TYPE B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;

```

RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=1436;
RA	Kirma N., Ferreira J.L., Baumstark B.R.;
RT	"Characterization of six type A strains of Clostridium botulinum that contain type B toxin gene sequences.";
RL	Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR	EML; AF295926; AA87132.1; -;
SQ	SEQUENCE 1291 AA; 150834 MW; D7CA07BAE2EB8CD2 CRC64;
Query Match	32.0%; Score 64; DB 2; Length 1291;
Best Local Similarity	62.5%; Pred. No. 1.6;
Matches 10; Conservative	5; Mismatches 1; Indels 0; Gaps 0;
QY	15 SLFNNFTVSEWLRVPK 30 : : : : : :
Dd	921 SMFLDFSVFWIRPK 936
RESULT 10	
Q933KO	PRELIMINARY; PRT; 1291 AA.
ID Q933KO	AC Q933KO;
DT 01-DEC-2001	(TrEMBLrel. 19, Created)
DT 01-DEC-2001	(TrEMBLrel. 19, Last sequence update)
DE TYPE-B CRYPTIC NEUROTOXIN.	
OS Clostridium botulinum.	
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;	
CC Clostridium.	
OX NCBI_TaxID=1491;	
[1]	
RN	SEQUENCE FROM N.A.
RC	STRAIN=593, AND 588;
RA	Kirma N., Ferreira J.L., Baumstark B.R.;
RT	"Characterization of six type A strains of Clostridium botulinum that contain type B toxin gene sequences.";
RL	Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR	EML; AF300466; AAL11499.1; -;
DR	EML; AF300465; AAL11498.1; -;
KW Neurotoxin.	
SQ SEQUENCE 1291 AA; 150843 MW; 7AC1737B0FA5A151 CRC64;	
Query Match	32.0%; Score 64; DB 2; Length 1291;
Best Local Similarity	62.5%; Pred. No. 1.6;
Matches 10; Conservative	5; Mismatches 1; Indels 0; Gaps 0;
QY	15 SLFNNFTVSEWLRVPK 30 : : : : : :
Dd	921 SMFLDFSVFWIRPK 936
RESULT 11	
Q45894	PRELIMINARY; PRT; 1296 AA.
ID Q45894	AC Q45894; P77780;
DT 01-NOV-1996	(TrEMBLrel. 01, Created)
DT 01-NOV-1996	(TrEMBLrel. 01, Last sequence update)
DE BONTOLINUM NEUROTOXIN TYPE A (TYPE A NEUROTOXIN).	
GZ BONT OR ATX.	
OS Clostridium botulinum.	
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;	
CC Clostridium.	
OX NCBI_TaxID=1491;	
[1]	
RN	SEQUENCE FROM N.A.
RC	STRAIN-KYOTO-F;
RX MEDLINE=94143603; PubMed=8310180;	
RA Williams A., East A.K., Lawson P.A., Collins M.D.;	
RT "Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A associated with infant botulism: comparison with other clostridial neurotoxins.";	
RT	


```

RX MEDLINE=98404102; PubMed=9732534;
RA East A.K., Bhandari M., Hiem S., Collins M.D.;
RT "Analysis of the botulinum neurotoxin type F gene clusters in
RT proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT barati.";
RL Curr. Microbiol. 37:262-268(1998).
DR EMBL: X81714; CAA57358.1; -.
DR EMBL: L35496; AAA23210.1; -.
DR EMBL: X70821; CAA50152.1; -.
DR EMBL: X99064; CAA67512.1; -.
DR HSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_Mpeptidse.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Neurotoxin.
SQ SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 30.5%; Score 61; DB 2; Length 1278;
Best Local Similarity 57.1%; Pred. No. 4.2;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPK 30
Db : ||::||::||::||

RESULT 13
Q90Y63 PRELIMINARY; PRT; 90 AA.
AC Q90Y63;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GONADOTROPIN-RELEASING HORMONE.
GN GNRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GNRH1 and GNRH2
RT precursors from bullfrog (Rana catesbeiana).";
RL J. Exp. Zool. 289:190-201(2001).
DR EMBL: AF188754; AAL05972.1; -.
SQ SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;

Query Match 29.0%; Score 58; DB 13; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db : |||||
26 HWSYGLRPG 34

RESULT 14
Q9PRH0 PRELIMINARY; PRT; 91 AA.
AC Q9PRH0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE GONADOLIBERIN PRECURSOR (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-
DE RH) (LULIBERIN).

```

```

OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguilloidei.
OC Anguillidae; Anguilla.
OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Okubo K., Suetake H., Aida K.;
RT "Expression of two gonadotropin-releasing hormone (GNRH) precursor
RT genes in various tissues of the Japanese eel and evolution of GNRH.";
RL Zool. Sci. 16:471-478(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Okubo K., Suetake H., Aida K.;
RT "A splicing variant for the prepro-mammalian gonadotropin-releasing
RT hormone (prepro-mGNRH) mRNA is present in the brain and various
RT peripheral tissues of the Japanese eel.";
RL Zool. Sci. 16:645-651(1999).
CC -|- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC EMBL: AB026989; BAA82608.1; -.
CC EMBL: AB026991; BAA83597.1; -.
CC InterPro: IPR002012; GNRH.
CC InterPro: IPR004079; Gonadoliberin1.
CC Pfam: PF00446; GNRH; 1.
CC PRINTS; PR01541; GONADOLIBRN1.
CC PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 32 MGNRH.
FT CHAIN 33 91 GNRH ASSOCIATED PEPTIDE.
SQ SEQUENCE 91 AA; 9893 MW; BA15C9DC08434A7B CRC64;

Query Match 29.0%; Score 58; DB 13; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db : |||||
24 HWSYGLRPG 32

RESULT 15
O92837 PRELIMINARY; PRT; 234 AA.
AC O92837;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VP3 (VP3 MINOR STRUCTURAL PROTEIN) (MINOR STRUCTURAL PROTEIN VP3)
DE (VIRAL COAT PROTEIN 3) (MINOR VIRAL COAT PROTEIN 3).
GN VP3.
OS Simian virus 40 (SV40).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K661;
RX MEDLINE=98216763; PubMed=9557685;
RA Lednický J.A., Arrington A.S., Stewart A.R., Cai X.M., Wong C.,
RA Jafar S., Murphy-Corb M., Butel J.S.;
RT "Natural isolates of simian virus 40 from immunocompromised monkeys
RT display extensive genetic heterogeneity: new implications for
RT polyomavirus disease.";
RL J. Virol. 72:3980-3990(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K661;
RA Lednický J.A., Arrington A.S., Stewart A.R., Dai X.M., Butel J.S.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: October 10, 2002, 15:58:41 ; Search time 4.82265 Seconds
(without alignments)
297.061 Million cell updates/sec
Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSGSLFNNFTVFWLRVVKVSASHLE 37
Scoring table:
Gapop 10.0 , Gapext 0.5
Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	57.0	1314	1	TETX_CLOTE
2	67	33.5	1296	1	BXG_CLOBO
3	65	32.5	1295	1	BXA1_CLOBO
4	63	31.5	1274	1	BXF_CLOBO
5	63	31.5	1290	1	BXE_CLOBO
6	62	31.0	1051	1	VP2_AHSV6
7	62	31.0	1295	1	BXA2_CLOBO
8	60	30.0	92	1	GON1_TUPGB
9	59	29.5	67	1	GON1_MACMU
10	59	29.5	92	1	GON1_HUMAN
11	58	28.0	61	1	GON1_SHEEP
12	58	29.0	63	1	GON1_MESAU
13	58	29.0	89	1	GON1_XENLA
14	58	29.0	90	1	GON1_MOUSE
15	58	29.0	91	1	GON1_PIG
16	58	29.0	92	1	GON1_RAT
17	57.5	28.7	352	1	COR2_SV40
18	56.5	28.2	94	1	GON1_HAPBU
19	56	28.0	92	1	GON1_CHICK
20	56	28.0	1250	1	BXE_CLOBO
21	56	28.0	1250	1	BXE_CLOBO
22	56	28.0	1290	1	BXC1_CLOBO
23	54.5	27.3	464	1	VNS5_TSWV1
24	54	27.0	10	1	GON1_ALLMI
25	53	26.5	99	1	GON1_DICLA
26	53	26.5	449	1	VNS5_INSVN
27	53	26.5	1196	1	BXCN_CLOBO
28	52	26.0	95	1	GON1_MORSA
29	52	26.0	95	1	GON1_PAGMA
30	52	26.0	95	1	GON1_SPAAU
31	52	26.0	467	1	VNS5_TSWVL
32	51.5	25.8	1539	1	SMCY_HUMAN
33	51.5	25.8	1548	1	SMCY_MOUSE

34	51.5	25.8	1554	1	SMCX_MOUSE	P41230 mus musculus
35	51.5	25.8	1560	1	SMCX_HUMAN	P41229 homo sapien
36	51	25.5	227	1	RHON_HUMAN	P52198 homo sapien
37	51	25.5	227	1	RHON_MOUSE	Q9qym5 mus musculus
38	50.5	25.2	1276	1	BXD_CLOBO	P19321 clostridium
39	50	25.0	90	1	GON8_RANDY	Q9iaa2 rana dybows
40	50	25.0	144	1	CNI_DROME	P49858 drosophila
41	50	25.0	585	1	YH70_SYNY3	P73627 synchocyst
42	49.5	24.8	344	1	COA2_POVJC	P03095 polyomaviru
43	49.5	24.8	1115	1	ITA3_DROME	O44386 drosophila
44	49	24.5	92	1	GON1_CAVPO	O54713 cavia porce
45	49	24.5	144	1	CNIH_HUMAN	O95406 homo sapien

ALIGNMENTS

RESULT 1
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).
OS Clostridium tetani.
OG Plasmid.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarasch W., Goretzki K., Henschen A., Engels J.,
Weller U., Hudek M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
homology with botulinum toxins";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CN3911;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE OF 742-1314 FROM N.A.
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
fragment C in Escherichia coli";
RL J. Bacteriol. 165:21-27(1986).
RN [4]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90210134; PubMed=2108021;
RA Krieglstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulphhydryl groups
in tetanus toxin";
RL Eur. J. Biochem. 188:39-45(1990).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
identification of cleavage sites";
RL Eur. J. Biochem. 202:41-51(1991).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396538;
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
neurotransmitter release and protease activity depend on zinc.";

```

RL EMBO J. 11:3577-3583(1992).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin.";
RN Nature 359:832-835(1992).
RL [8]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX MEDLINE=97475217; PubMed=9334741;
RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA Sax M.;
RT "Structure of the receptor binding fragment HC of tetanus
RT neurotoxin.";
RL Nat. Struct. Biol. 4:788-792(1997).
CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
CC BOND OF SYNAPTOSREVIN-2.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-|-PHE-77 BOND IN
CC SYNAPTOSREVIN.
CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC AND ARE NON-TOXIC AFTER SEPARATION.
CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC GANGLIOSIDE RECEPTORS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04436; CAA28033.1; -
DR EMBL; M12739; AAA23282.1; -
DR EMBL; X06214; CAA29564.1; -
DR PIR; A25689; BTCLTN.
DR PDB; 1AF9; 29-APR-98.
DR PDB; 1A8D; 14-OCT-98.
DR MEROPS; M27.001; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
KW 3D-structure.
FT INIT_MET 0
FT CHAIN 1 456
FT CHAIN 457 1314
FT CHAIN 1314
FT METAL 232 232
FT METAL 233 233
FT ACT_SITE 233 233
FT METAL 236 236
FT TRANSMEM 226 246
FT TRANSMEM 669 689
FT DISULFID 438 466
FT DISULFID 1076 1092
FT SEQUENCE 1314 AA; 150550 MW; 134C3657133FE81D CRC64;
Query Match 57.08; Score 114; DB 1; Length 1314;
Best Local Similarity 95.58; Pred. No. 3.9e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 945 MFNFTVSWFLRVPKVSASHLE 966
:|||||
RESULT 2
BXG_CLOBO
ID ID BXG_CLOBO STANDARD; PRT; 1296 AA.
AC Q60393;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BONT/G)
DE (Bontoxilysin G)
GN BOTG.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium
OC Clostridium
OX NCBI_Taxid=1491;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=113 / 30;
RX MEDLINE=94092745; PubMed=8268233;
RA Campbell K., Collins M.D., East A.K.;
RT "Nucleotide sequence of the gene coding for Clostridium botulinum
RT (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT with other clostridial neurotoxins.";
RL Biochim. Biophys. Acta 1216:487-491(1993).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74162; CAA52275.1; -
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 1 441
FT CHAIN 442 1296
FT METAL 229 229
FT ACT_SITE 230 230
FT METAL 233 233
FT DISULFID 435 449
FT SEQUENCE 1296 AA; 149013 MW; DC8E47E15F665C31 CRC64;
Query Match 33.58; Score 67; DB 1; Length 1296;

```


heavy chain (H).
-!- SUBCELLULAR LOCATION: Secreted.
-!- PHARMACEUTICAL: Available under the name BOTOX(R) (Allergan) for the treatment of strabismus and blepharospasm associated with dystonia and cervical dystonia. Also used for the treatment of hemifacial spasm and a number of other neurological disorders characterized by abnormal muscle contraction.
-!- MISCELLANEOUS: There are seven antigenically distinct forms of botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X52066; CAA36289.1; -
DR EMBL; M30196; AAA23262.1; -
DR EMBL; X92973; CAA63551.1; -
DR EMBL; D67030; BAA11051.1; -
DR EMBL; M27892; AAA23269.1; -
DR PIR; A35294; BTCLAB.
DR PIR; S09492; S09492.
DR PDB; 3BTA; 01-OCT-99.
DR MEROPS; M27.002; -
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_Mpeptidase.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR PRODom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW Pharmaceutical; 3D-structure.
FT INIT_MET 0 0
FT CHAIN 1 447
FT CHAIN 448 1295
FT METAL 222 222
FT METAL 223 223
FT ACT_SITE 226 226
FT METAL 261 261
FT METAL 261 261
FT DISULFID 429 453
FT DISULFID 1234 1279
FT TRANSMEM 626 646
FT TRANSMEM 655 675
FT VARIANT 26 26
FT MUTAGEN 261 261
FT MUTAGEN 265 265
FT MUTAGEN 365 365
FT CONFLICT 1 1
FT CONFLICT 479 479
FT CONFLICT 875 875
FT CONFLICT 891 891
SQ SEQUENCE 1295 AA; 149322 MW; 858342F754862579 CRC64;
Query Match 32.5%; Score 65; DB 1; Length 1295;
Best Local Similarity 56.2%; Pred. No. 0.36;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Qy 15 SLFNFTVSWFLVRPK 30
I:: I:: I:: I:: I::
Db 935 SWYENFTSWFIRPK 950
RESULT 4
BXF_CLOBO *
ID BXF_CLOBO STANDARD; PRT; 1274 AA.
AC P30996;
DC 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BONT/F)
DE (Bontoxilysin F).
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23387;
RX MEDLINE=93012902; PubMed=1398040;
RA East A.K., Richardson P.T., Alloway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.;
RT "Sequence of the gene encoding type F neurotoxin of Clostridium botulinum";
RL FEMS Microbiol. Lett. 75:225-230(1992).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC STRAIN=HOBBS FT10;
RX MEDLINE=94297488; PubMed=7764998;
RA East A.K., Collins M.D.;
RT "Conserved structure of genes encoding components of botulinum neurotoxin complex M and the sequence of the gene coding for the nontoxic component in nonproteolytic Clostridium botulinum type F";
RL Curr. Microbiol. 29:69-77(1994).
RN [3]
RP SEQUENCE OF 634-1002 FROM N.A.
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and specific identification of neurotoxin types B, E, and F";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94230352; PubMed=8175689;
RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA Roques B., Fyke E.M., Suedhof T.C., Jahn R., Niemann H.;
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and F botulinum neurotoxins and tetanus toxin";
RL J. Biol. Chem. 269:12764-12772(1994).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-1-LYS-59 BOND OF SYNAPTOSOMES-1 AND -2.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No detected action on small molecule substrates.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY, WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M92906; AAA23263.1; -
DR EMBL; S73676; AAC60475.1; -
DR EMBL; X70820; CAA50151.1; -
DR EMBL; X70816; CAA50147.1; -
DR HSSP; P10845; 3BTA.

```

DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_Mtpeptidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PROSITE; PS001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT CHAIN 1 436
FT CHAIN 437 1274
FT ACT_SITE 227 228
FT METAL 231 231
FT DISULFID 429 445
SQ SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match 31.5%; Score 63; DB 1; Length 1274;
Best Local Similarity 56.2%; Pred. No. 0.68;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNFTVSFWLRVPK 30
DB 928 SRYQNFISFWVRIPK 943
I : ||:||||:|
| : ||:||||:|

RESULT 5
BXB_CLOBO
ID BXB_CLOBO STANDARD; PRT; 1290 AA.
AC P10844; P10843;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BONT/B)
DE (Bontoxilysin B).
GN BOTB.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_taxid=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92384550; PubMed=1514783;
RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA Minton N.P.;
RT "Molecular cloning of the Clostridium botulinum structural gene
RT encoding the type B neurotoxin and determination of its entire
RT nucleotide sequence."
RL Appl. Environ. Microbiol. 58:2345-2354(1992).
RN [2]
RP SEQUENCE OF 35-245 FROM N.A.
RC STRAIN=NCYC 7273;
RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 633-993 FROM N.A.
RC STRAIN=NCYC 7273;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F."
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-44 AND 441-466.
RC STRAIN=657;
RX MEDLINE=89000987; PubMed=3139097;
RA Dasgupta B.R., Datta A.;
RT "Botulinum neurotoxin type B (strain 657): partial sequence and
RT similarity with tetanus toxin."
RL Biochimie 70:811-817(1988).
RN [5]
RP SEQUENCE OF 1-16 AND 441-458.
RC STRAIN=OKRA;
RX MEDLINE=85197963; PubMed=3888113;

```

```

RA Schmidt J.J., Sathyamoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequences of botulinum neurotoxins types B and
RT E."
RN Arch. Biochem. Biophys. 238:544-548(1985).
RX [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93054694; PubMed=1429690;
RA Schlavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
RT "Botulinum neurotoxins are zinc proteins."
RN J. Biol. Chem. 267:23479-23483(1992).
RX [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schlavo G., Benfenati F., Foulain B., Rossetto O., de Laureto P.P.,
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin."
RL Nature 359:832-835(1992).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
CC SYNAPTOSOMAL-2.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81186; AAA23211.1; -
CC EMBL; Z11934; CAA77991.1; -
CC EMBL; X70817; CAA50148.1; -
CC PIR; S07128; S07128.
CC PIR; S07155; S07155.
CC PIR; S08562; S08562.
CC PIR; S08573; S08573.
CC PIR; S08574; S08574.
CC PIR; A48940; A48940.
CC HSP; P10845; 3BTA.
CC MEROPS; M27.002; -.
CC InterPro: IPR000395; Bontoxilysin.
CC InterPro: IPR000130; Zn_Mtpeptidase.
CC Pfam; PF01742; Peptidase_M27; 1.
CC PRINTS; PR00760; BONTOXILYSIN.
CC PRODOM; PD001963; Bontoxilysin; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0
FT CHAIN 1 440
FT CHAIN 441 1290
FT METAL 229 229
FT ACT_SITE 230 230
FT METAL 233 233
FT DISULFID 436 445
FT CONFLICT 29 29
FT CONFLICT 217 217
FT CONFLICT 224 224

```

```
FT CONFLICT 463 463 S -> R (IN REF. 4).
SQ SEQUENCE 1290 AA; 150670 MW; D21745E2C024DF43 CRC64;

Query Match 31.5%; Score 63; DB 1; Length 1290;
Best Local Similarity 62.5%; Pred. No. 0.69;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
Db 920 SVFLDSVSWIRIPK 935
I:I :I:|I|:|I|I
I:I :I:|I|:|I|I

RESULT 6
VP2_AHSV6 STANDARD; PRT; 1051 AA.
AC VP21024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Outer capsid protein VP2.
GN S2 OR L2.
OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
OS (serotype 6)).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=86060;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98278331; PubMed=9617769;
RA Williams C.F., Inoue T., Lucas A.M., Zanotto P., Roy P.;
RT "The complete sequence of four major structural proteins of African
RT horse sickness virus serotype 6: evolutionary relationships within
RT and between the orbiviruses.";
RL Virus Res. 53:53-73(1998).
CC -!- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
CC WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
CC MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF021235; AAC40894.1; -
CC InterPro; IPR001742; Orbl_VP2.
CC Pfam; PF00898; Orbl_VP2; 2.
CC Coar protein.
SQ SEQUENCE 1051 AA; 122326 MW; 2B04DB9E389F4B5F CRC64;

Query Match 31.0%; Score 62; DB 1; Length 1051;
Best Local Similarity 40.7%; Pred. No. 0.76;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 11 SSGPSLNNFTVSFWLRVPKVSASHLE 37
Db 630 TEGVTFSKRFVSFWYRVEKITKHL 656
I:I :I:|I|:|I|I
I:I :I:|I|:|I|I

RESULT 7
BXA2_CLOBO STANDARD; PRT; 1295 AA.
AC Q45894; P77780;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
DE chain; Botulinum neurotoxin A, heavy-chain].
DE BOTA OR BNA OR ATX.
GN Clostridium botulinum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Kyoto-F;
RC MEDLINE=94143603; PubMed=8310180;
RA Williams A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium
RT botulinum type A associated with infant botulism: comparison with
RT other clostridial neurotoxins.";
RL Res. Microbiol. 144:547-556(1993).
RN [2]
RP SEQUENCE OF 1-65 FROM N.A.
RX STRAIN=Kyoto-F;
RC MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC -!- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC binds with high affinity to peripheral neuronal presynaptic
CC membrane, is then internalized by receptor-mediated endocytosis.
CC The C-terminus of the heavy chain (H) is responsible for the
CC adherence of the toxin to the cell surface while the N-terminus
CC mediates transport of the light chain from the endocytic vesicle
CC to the cytosol. After translocation, the light chain (L)
CC hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking
CC neurotransmitter release. Inhibition of acetylcholine release
CC results in flaccid paralysis, with frequent heart or respiratory
CC failure (by similarity).
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H) (by similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X73423; CAA51824.1; -
CC EMBL; X87974; CAA61234.1; -
CC HSP; P10845; 3B7A.
CC MEROPS: M27.002; -.
CC InterPro; IPR000395; Bontoxilysin.
CC Pfam; PF01742; Peptidase_M27; 1.
CC PRINTS; PR00760; BONTOXILYSIN.
CC ProDom; PD001963; Bontoxilysin; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0 BY SIMILARITY.
FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 223 223 BY SIMILARITY.
FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 429 453 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1234 1279 BY SIMILARITY.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675 POTENTIAL.
SQ SEQUENCE 1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;
```



```
RESULT 10
GONL_HUMAN
ID GONL_HUMAN STANDARD; PRT; 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
DE peptide I].
GN GNRH1 OR GNRH OR LHRH.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89366682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RT hormone gene.";
RL Nucleic Acids Res. 17:6403-6403(1989).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone.";
RL Nature 311:666-668(1984).
[4]
RN SEQUENCE OF 24-33.
RX MEDLINE=83126573; PubMed=6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
RT biosynthesized in the human placenta.";
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC Lutrepulse or Lutrelef (Perring Pharmaceuticals) and Relisorm
CC (Serono).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X01059; CAA25526.1; -
CC EMBL; M12578; AAA35916.1; -
CC EMBL; X15215; CAA33285.1; -
CC PIR; A01410; RHUG.
CC PIR; A26173; A26173.
CC PIR; S05308; S05308.
CC MIM; 152760; -
CC InterPro: IPR002012; GNRH.
CC InterPro: IPR004079; GonadoliberinI.
CC Pfam: PF00446; GNRH; 1.
CC PRINTS; PK01541; GONADOLIBRNI.

DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal.
FT SIGNAL 1 23 PROGONADOLIBERIN I.
FT CHAIN 24 92 GONADOLIBERIN I.
FT PEPTIDE 24 33 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 37 92 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 26 26 ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT CONFLICT 16 16 W -> S (IN REF. 3).
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 29.5%; Score 59; DB 1; Length 92;
Best Local Similarity 52.4%; Pred. No. 0.13;
Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLFNNF 20
DB 25 HWSYGLRPGGKRAENLIDSF 45
||||||| :|:|
| | | | | |

RESULT 11
GONL_SHEEP
ID GONL_SHEEP STANDARD; PRT; 61 AA.
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
[1]
RN SEQUENCE OF 12-61 FROM N.A.
RX MEDLINE=72094314; PubMed=4550508;
RA Burgus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vale W., Guillemin R.;
RT "Primary structure of the ovine hypothalamic luteinizing hormone-
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-Edman degradation).";
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U02517; AAA03433.1; -
CC PIR; A93780; RSHSG.
CC InterPro: IPR002012; GNRH.
CC Pfam: PF00446; GNRH; 1.
CC PROSITE; PS00473; GNRH; 1.
```

KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >61 PRONADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6828 MW; 63962AAAE319B8F0 CRC64;

Query Match 29.0%; Score 58; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSYGLRPG 10
|||||

RESULT 12
GON1_MESAU
ID GON1_MESAU STANDARD; PRT; 63 AA.
AC 009163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
(Luteinizing hormone releasing hormone I) (Gonadotropin releasing
hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
GN GNRH OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TAXID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS: IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U91938; AAB51302.1; -
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1 PRONADOLIBERIN I.
FT CHAIN 1 >63 GONADOLIBERIN I.
FT PEPTIDE 1 10
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY
SIMILARITY).
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY (BY SIMILARITY).
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
SIMILARITY).
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY

FT NON_TER 63 63 SIMILARITY).
SQ SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;

Query Match 29.0%; Score 58; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 2 HWSYGLRPG 10
|||||

RESULT 13
GON1_XENLA
ID GON1_XENLA STANDARD; PRT; 89 AA.
AC P45656;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Gonadoliberein I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
(LH-RH) (Luliberin I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TAXID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Hayes W.P., Wray S., Battey J.F.;
RX MEDLINE=94185563; PubMed=8137750;
RT "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a
mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L28040; AAA49728.1; -
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23 PRONADOLIBERIN I.
FT CHAIN 24 89 GONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
FT PEPTIDE 37 89 PEPTIDE.
FT PEPTIDE 37 85 GNRH-ASSOCIATED PEPTIDE I (GAP).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;

Query Match 29.0%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||

```
Db      25 HWSYGLRPG 33
RESULT 14
GONL_MOUSE STANDARD; PRT; 90 AA.
AC      P13562;
DT      01-JAN-1990 (Rel. 13, Created)
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE      (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE      hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE      I].
GN      GNRH1 OR GNRH.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87069928; PubMed=3024317;
RA      Mason A.J., Haylick J.S., Zoeller R.T., Young W.S. III,
RA      Phillips H.S., Nikolics K., Seeburg P.H.;
RT      "A deletion truncating the gonadotropin-releasing hormone gene is
RT      responsible for hypogonadism in the hpg mouse.";
RL      Science 234:1366-1371(1986).
CC      -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC      THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC      HORMONES.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M14872; AAA37717.1; -.
DR      MGD; MGI:95789; GnRH.
DR      InterPro: IPR002012; GnRH.
DR      InterPro: IPR004079; GonadoliberinI.
DR      Pfam; PF00446; GnRH; 1.
DR      PRINTS; PR01541; GONADOLIBRN1.
DR      PROSITE; PS00473; GNRH; 1.
KW      Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW      Placenta; Signal.
FT      SIGNAL 1 21
FT      CHAIN 22 90 PROGONADOLIBERIN I.
FT      PEPTIDE 22 31 GONADOLIBERIN I.
FT      PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT      ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT      ACTIVITY.
FT      MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT      MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
SQ      SEQUENCE 90 AA; 10337 MW; 1C0766FA826E4D9 CRC64;
Query Match 29.0%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2 HWSYGLRPG 10
DB      23 HWSYGLRPG 31
RESULT 15
GONL_PIG STANDARD; PRT; 91 AA.
AC      P49921;
DT      01-JAN-1990 (Rel. 13, Created)
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE      (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE      hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE      I].
GN      GNRH1 OR GNRH.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9823;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=72065376; PubMed=4946067;
RA      Baba Y., Arimura A., Schally A.V.;
RT      "Structure of the porcine LH- and FSH-releasing hormone. II.
RT      Confirmation of the proposed structure by conventional sequential
RT      analyses.";
RL      Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN      [3]
RP      SYNTHESIS OF GONADOLIBERIN.
RX      MEDLINE=72065376; PubMed=4942726;
RA      Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT      "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RT      phase method.";
RL      Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN      [4]
RP      SYNTHESIS OF GONADOLIBERIN.
RX      MEDLINE=72117544; PubMed=4946275;
RA      Baba Y., Arimura A., Schally A.V.;
RT      "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
RL      Biochem. Biophys. Res. Commun. 45:483-487(1971).
CC      -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC      THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC      HORMONES.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; L32864; AAA31066.1; -.
DR      PIR; A01411; RHPGG.
DR      InterPro: IPR002012; GnRH.
DR      InterPro: IPR004079; GonadoliberinI.
DR      Pfam; PF00446; GnRH; 1.
DR      PRINTS; PR01541; GONADOLIBRN1.
DR      PROSITE; PS00473; GNRH; 1.
KW      Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW      Placenta; Signal.
FT      SIGNAL 1 23 PROGONADOLIBERIN I.
FT      CHAIN 24 91 GONADOLIBERIN I.
FT      PEPTIDE 24 33 GONADOLIBERIN I.
FT      PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
FT      ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT      ACTIVITY.
FT      MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT      MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ      SEQUENCE 91 AA; 10090 MW; 8340474F3DDAA99 CRC64;
Query Match 29.0%; Score 58; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 2 HWSYGLRPG 10
| | | | | | | | | |
Db 25 HWSYGLRPG 33

Search completed: October 10, 2002, 16:06:33
Job time : 5.82265 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:00:11 ; Search time 8.61752 Seconds

(without alignments)

412.567 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHWSYGLRPGSSGSLFNNFTVFWLRVPKVSASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	57.0	1315	1 BTCLTN	tentoxylisin (EC 3
2	67	33.5	1297	2 S39791	neurotoxin - Clost
3	65	32.5	1296	1 BTCLAB	bontoxylisin (EC 3
4	64	32.0	1268	2 S33411	botulinum neurotox
5	64	32.0	1291	2 I40631	non-proteolytic bo
6	63	31.5	369	2 S48109	neurotoxin type F
7	63	31.5	1274	2 I40813	neurotoxin type F
8	63	31.5	1291	1 A48940	bontoxylisin (EC 3
9	62	31.0	1296	2 I40645	botulinum neurotox
10	61	30.5	366	2 S48110	neurotoxin type F
11	59	29.5	67	2 I78541	gonadoliberin prec
12	59	29.5	92	1 RHUG	gonadoliberin prec
13	58	29.0	10	1 RHGG	gonadoliberin - pi
14	58	29.0	10	1 RHSG	gonadoliberin - sh
15	58	29.0	89	2 I51423	gonadoliberin prec
16	58	29.0	90	1 RHMSG	gonadoliberin prec
17	58	29.0	92	1 RHRTG	gonadoliberin prec
18	57.5	28.7	352	1 VVVP24	coat protein vp2 -
19	56.5	28.2	98	2 I50739	gonadotropin-relea
20	56	28.0	92	2 I50644	gonadoliberin I pr
21	56	28.0	367	2 S48106	neurotoxin type E
22	56	28.0	1251	2 JH0256	botulinum neurotox
23	56	28.0	1252	2 S21178	botulinum neurotox
24	56	28.0	1291	2 S45431	botulinum neurotox
25	56	28.0	1291	2 A43777	botulinum neurotox
26	54.5	27.3	91	2 JC7393	medaka-type gonad
27	54.5	27.3	464	1 MNVUWC	nonstructural prot
28	54	27.0	10	1 RHAQ1	gonadoliberin I -
29	53.5	26.8	251	2 AD1669	3'-exo-deoxyribonu

30	53	26.5	449	2 S23158	nucleocapsid prote
31	53	26.5	1196	2 JQ1467	toxin, nontoxic co
32	53	26.5	1196	2 S46430	botulinum neurotox
33	53	26.5	1285	2 S70582	botulinum neurotox
34	52.5	26.2	251	2 AF1297	3'-exo-deoxyribonu
35	52.5	26.2	1999	2 AB2018	hypothetical prote
36	52	26.0	210	2 T18703	hypothetical prote
37	52	26.0	467	1 MNVUW1	nonstructural prot
38	52	26.0	916	2 T04752	aspartate kinase (
39	51.5	25.8	564	2 I48776	spermatogenesis re
40	51.5	25.8	1033	2 I48775	SmcX protein (esca
41	51.5	25.8	1560	2 I54361	SMCX protein - hum
42	51	25.5	436	2 D84782	probable proline t
43	51	25.5	637	2 F90257	hypothetical prote
44	51	25.5	1261	2 AI3471	5-methyltetrahydro
45	50.5	25.2	230	2 T27498	hypothetical prote

ALIGNMENTS

RESULT 1

BTCLTN
tentoxylisin (EC 3.4.24.68) precursor - Clostridium tetani
N:Alternate names: tetanus neurotoxin
C:Species: Clostridium tetani
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,
EMBO J. 5, 2495-2502, 1986
A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b
A:Reference number: A25689; MUID:87053814
A:Accession: A25689
A:Molecule type: DNA
A:Residues: 1-1315 <EIS>
A:Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770
R:Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A:Title: The complete nucleotide sequence of tetanus toxin.
A:Reference number: A25757; MUID:87040747
A:Accession: A25757
A:Molecule type: DNA
A:Residues: 1-1315 <FAI>
A:Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
A:Experimental source: strain CN3911
R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C i
A:Reference number: A25194; MUID:86085672
A:Accession: A25194
A:Molecule type: DNA
A:Residues: 743-1315 <FA2>
A:Cross-references: GB:M12739; NID:g14920; PIDN:AAA23282.1; PID:g14921
A:Accession: B25194
A:Molecule type: protein
A:Residues: 865-894 <FA3>
R:Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termin
A:Reference number: A60759; MUID:90035436
A:Accession: A60759
A:Molecule type: protein
A:Residues: 461-475 <MAT>
R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A:Reference number: JS0098; MUID:89093918
A:Contents: annotation: epitope region
R:Schlavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,
Nature 359, 832-835, 1992
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A:Reference number: S27125; MUID:93063293
A:Contents: annotation

A:Residues: 1-1296 <BN>
A:Cross-references: GB:M30196; NID:g144864; PIDN:AAA23262.1; PID:g144865
A:Experimental source: strain 62A, subtype A
A:R:Thompson, D.E.; Brehm, J.K.; Cultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson, E.
Eur. J. Biochem. 189, 73-81, 1990
A:Title: The complete amino acid sequence of the Clostridium botulinum type A neurotoxin
A:Reference number: S09492; MUID:90235864
A:Accession: S09492
A:Molecule type: DNA
A:Residues: 1', 3'-26, 'V', 28-1296 <THO>
A:Cross-references: EMBL:X52066; NID:g40381; PIDN:CAA36289.1; PID:g40382
A:Experimental source: NCRC 2916
A:R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376, 41-44, 1995
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin components
A:Reference number: S67988; MUID:96096783
A:Accession: S68220
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-12 <FUJ>
A:Cross-references: EMBL:D67030; DDBJ:D50421; NID:g2160224
A:R:Betley, M.J.; Somers, E.; DasGupta, B.R.
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-terminal region
A:Reference number: A33401; MUID:89350959

A;Reference number: A33401; MUID:89350959
A;Accession: A33401
A;Molecule type: DNA

A; Cross-references: GB:M27892; NID:g144880; PIDN:AAA23269.1; PID:g551776
R; Gimenez, J.A.; DasGupta, B.R.

J. Protein Chem. 12, 351-363, 1993

A;Accession: A53884

A;Status: preliminary
A;Molecule type: protein
A;Residues: 867-880;1148-1217, 'Y', 1219 <GIM>
A;Experimental source: strain Hall
A;Note: sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after R. DasGupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A;Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and a
A;Reference number: A60025; MUID:91120847
A;Accession: A60025
A;Molecule type: protein
A;Residues: 2-6;445-453, 'X', 455-457 <DASI>
R;DasGupta, B.R.; Foley, J.; Niece, R.
Biochemistry 26, 4162, 1987
A;Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A;Reference number: A27000

A; Molecule type: protein
A; Residues: 2-47 <DAS2>
R; Binz, T.; Blasi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.

A;Title: Proteolysis of SNAP-25 by types E and A botulinal neurotoxins.
A;Reference number: A49708; MUID:94124495

A;Contents: annotation

A;Gene: atx; bota

C;Function: Query Match 32.5% Score 65; DB 1; Length 1296;
A;Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptoosomal-associated vesicles
C;Superfamily: tetanus toxin
C;Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane
F;2-444/Product: bontoxilysin A light chain #status experimental <LGT>
F;445-1296/Product: bontoxilysin A heavy chain #status experimental <HVT>
F;223,227/Binding site: zinc (His) #status predicted
F;224/Active site: Glu #status predicted
Best Local Similarity 56.2%; Pred. No. 1;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;


```
QY 15 SLFNNFTVSFWLRVPK 30
      1: 11:111:111
Db 936 SMYENFTSFWIRIPK 951
      1: 11:111:111

RESULT 4
S3411
botulinum neurotoxin type F - Clostridium barati
C:Species: Clostridium barati
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.
FEMS Microbiol. Lett. 108, 175-182, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin
A:Reference number: S33411; MUID:93252228
A:Accession: S33411
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1268 <THO>
A:Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139
A:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 32.0%; Score 64; DB 2; Length 1268;
Best Local Similarity 62.5%; Pred. No. 1.4;
Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
      1: 11:111:111
Db 920 SRYQNFVSFWIRIPK 935
      1: 11:111:111

RESULT 5
I40631
non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40631; S48103; S48104; S36015
R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Curr. Microbiol. 28, 101-110, 1994
A:Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botulinum
A:Reference number: I40631; MUID:94122659
A:Accession: I40631
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1291 <RES>
A:Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
A:Reference number: S48103; MUID:94013372
A:Accession: S48103
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 634-843, 'T', 845, 'N', 847-994 <CAM2>
A:Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
A:Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)
C:Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic s
C:Genetics:
A:Gene: bont/b
C:Superfamily: tetanus toxin
C:Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LIGHT>
F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVV>
F:230,234/Binding site: zinc (HIS) #status predicted

F:231/Active site: Glu #status predicted

Query Match 32.0%; Score 64; DB 2; Length 1291;
Best Local Similarity 62.5%; Pred. No. 1.4;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
      1: 11:111:111
Db 921 SMFLDFSFWIRIPK 936
      1: 11:111:111

RESULT 6
S48109
neurotoxin type F - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
C:Accession: S48109
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48109
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-369 <CAM>
A:Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C:Superfamily: tetanus toxin

Query Match 31.5%; Score 63; DB 2; Length 369;
Best Local Similarity 56.2%; Pred. No. 0.48;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
      1: 11:111:111
Db 295 SRYQNFVSFWIRIPK 310
      1: 11:111:111

RESULT 7
I40813
neurotoxin type F - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40813; S48108
R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson,
FEMS Microbiol. Lett. 96, 225-230, 1992
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A:Reference number: I40644
A:Accession: I40813
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1274 <RES>
A:Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48108
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 634-1002 <CAM>
A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 31.5%; Score 63; DB 2; Length 1274;
Best Local Similarity 56.2%; Pred. No. 2;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
      1: 11:111:111
Db 928 SRYQNFVSFWIRIPK 943
      1: 11:111:111
```

A48940
bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
N; Alternate names: botulinum neurotoxin type B (BoNT/B)
C; Species: Clostridium botulinum
C; Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
C; Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
R; Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A; Title: Molecular cloning of the Clostridium botulinum structural gene encoding the type B
A; Reference number: A48940; MUID:92384550
A; Accession: A48940
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-1291 <WHB>
A; Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
A; Experimental source: type B, Danish
A; Note: sequence extracted from NCBI backbone (NCBI:112080, NCBI:P:112081); this publica
R; Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A; Title: Gene probes for identification of the botulin neurotoxin gene and specific id
A; Reference number: S48103; MUID:94013372
A; Accession: S48105
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 634-994 <CAM>
A; Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
A; Experimental source: proteolytic type B, strain NCTC 7273
R; Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL data library, April 1992
A; Description: Partial amino acid sequence of botulinum neurotoxin type B and comparison
A; Reference number: S21575
A; Accession: S21575
A; Molecule type: DNA
A; Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZA>
A; Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384
R; Kurazono, H.; Wochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K.
J. Biol. Chem. 267, 14721-14729, 1992
A; Title: Minimal essential domains specifying toxicity of the light chains of tetanus to
A; Reference number: A42871; MUID:92340509
A; Accession: A42871
A; Status: nucleic acid sequence not shown
A; Molecule type: mRNA
A; Residues: 1-313, 'S', 315-451 <KUR>
A; Experimental source: strain Okra
A; Note: sequence extracted from NCBI backbone (NCBI:P:109365)
R; Dasgupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A; Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with
A; Reference number: S07155; MUID:8900987
A; Accession: S07155
A; Molecule type: protein
A; Residues: 2-29, 'M', 31-45 <DAS>
A; Accession: S08562
A; Molecule type: protein
A; Residues: 442-463, 'R', 465-467 <DA2>
R; Schmidt, J.J.; Sathyanarayana, V.; Dasgupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A; Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A; Reference number: S07128; MUID:85197963
A; Accession: S07128
A; Status: preliminary
A; Molecule type: protein
A; Residues: 2-16 <SCH1>
A; Accession: S08573
A; Status: preliminary
A; Molecule type: protein
A; Residues: 2-17 <SCH2>
A; Accession: S08574
A; Status: preliminary
A; Molecule type: protein
A; Residues: 442-459 <SCH3>

Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPK 30
: |||:||||:|

Db 297 YQNFISFWVRIPK 310

RESULT 11

I78541

gonadoliberin precursor - rhesus macaque (fragment)

N;Alternate names: luteinizing hormone releasing hormone

C;Species: Macaca mulatta (rhesus macaque)

C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999

C;Accession: I78541

R;Ma, Y.J.; Costa, M.E.; Ojeda, S.R.

Neuroendocrinology 60, 346-359, 1994

A;Title: Developmental expression of the genes encoding transforming growth factor alpha

A;Reference number: I58134; MUID:95124501

A;Accession: I78541

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-67 <RES>

C;Cross-references: GB:S75918; NID:g912831; PIDN:AAB33096.1; PID:g912832

C;Superfamily: gonadoliberin

Query Match

Best Local Similarity 29.5%; Score 59; DB 2; Length 67;

Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLFNNF 20

|||||

Db 7 HWSYGLRPGGKRDAENLWDSF 27

RESULT 12

RHHUG

gonadoliberin precursor [validated] - human

N;Alternate names: gonadotropin releasing hormone (GNRH); luteinizing hormone releasing

C;Species: Homo sapiens (man)

C;Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000

C;Accession: S05308; A26173; A93342; A90108; A01410; S45718

R;Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.

Nucleic Acids Res. 17, 6403-6404, 1989

A;Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone gene

A;Reference number: S05308; MUID:89366682

A;Accession: S05308

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-92 <HAY>

A;Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956

R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.

Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986

A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot

A;Reference number: A94090; MUID:86034338

A;Accession: A26173

A;Molecule type: mRNA

A;Residues: 1-92 <ADE>

A;Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749

A;Experimental source: hypothalamus

R;Seeburg, P.H.; Adelman, J.P.

Nature 311, 666-668, 1984

A;Title: Characterization of cDNA for precursor of human luteinizing hormone releasing h

A;Reference number: A93342; MUID:85012739

A;Accession: A93342

A;Molecule type: mRNA

A;Residues: 1-15,'S',17-92 <SEE>

A;Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357

A;Experimental source: placenta

R;Tan, L.; Rousseau, P.

Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982

A;Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in

A;Reference number: A90108; MUID:83126573

A;Accession: A90108

A;Molecule type: protein

A;Residues: 24-33 <TAN>

A;Experimental source: placental trophoblasts

R;Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amste

FEBS Lett. 346, 203-206, 1994

A;Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by

A;Reference number: S45718; MUID:94283597

A;Contents: annotation; degradation pathway of synthetic hormone

C;Genetics:

A;Gene: GDB:GNRH; LHRH; GRH

A;Cross-references: GDB:133746; OMIM:227200; OMIM:152760

A;Map position: 8p21-8p11.2

A;Introns: 47/3; 79/3

C;Function:

A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitrop

A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activi

C;Superfamily: gonadoliberin

C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-92/Product: progadoliberin #status predicted <PGN>

F;24-33/Product: gonadoliberin #status experimental <MAT>

F;37-92/Product: gonadoliberin-associated protein #status predicted <GAP>

F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi

F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match

Best Local Similarity 29.5%; Score 59; DB 1; Length 92;

Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLFNNF 20

|||||

Db 25 HWSYGLRPGGKRDAENLWDSF 45

RESULT 13

RHPGG

gonadoliberin - pig

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997

C;Accession: A01411

R;Baba, Y.; Matsuo, H.; Schally, A.V.

Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A;Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A;Reference number: A90172; MUID:72114303

A;Accession: A01411

A;Molecule type: protein

A;Residues: 1-10 <BAB>

R;Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A;Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A;Reference number: A90176; MUID:72065376

A;Contents: annotation; synthesis

A;Note: the synthetic and natural hormones have the same physicochemical and biologic

R;Baba, Y.; Arimura, A.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A;Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A;Reference number: A90175; MUID:72117544

A;Contents: annotation

A;Note: Trp-3 appears to be essential for biological activity

C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

C;Superfamily: gonadoliberin

C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match

Best Local Similarity 29.0%; Score 58; DB 1; Length 10;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

|||||

Db 2 HWSYGLRPG 10

```
RESULT 14
RHSBG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; Bl
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor
A:Reference number: A93780; MUID:72094314
A:Accession: A93780
A:Molecule type: protein
A:Residues: 1-10 <BUR>
A>Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 29.0%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10

RESULT 15
IS1423
gonadoliberin precursor - African clawed frog
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: IS1423
R:Hayes, W.P.; Wray, S.; Battey, J.F.
Endocrinology 134, 1835-1845, 1994
A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved doma
A:Reference number: IS1423; MUID:94185563
A:Accession: IS1423
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <HAY>
A:Cross-references: GB:L28040; NID:9496291; PIDN:AAA49728.1; PID:9496292
C:Genetics:
A:Gene: GnRH-I
C:Superfamily: gonadoliberin

Query Match 29.0%; Score 58; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
```

Search completed: October 10, 2002, 16:12:10
Job time : 9.69444 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 7.19231 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSGPSLKLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB ID	Description
1	106	60.9	40	2	US-08-460-502-10	Sequence 10, Appl
2	101	58.0	40	2	US-08-460-502-11	Sequence 11, Appl
3	73.5	42.2	34	5	PCT-US95-13841-28	Sequence 28, Appl
4	73.5	42.2	43	5	PCT-US95-13841-26	Sequence 26, Appl
5	73	42.0	42	1	US-08-446-692-20	Sequence 20, Appl
6	73	42.0	42	2	US-08-488-351A-20	Sequence 20, Appl
7	72	41.4	15	3	US-09-100-414B-1	Sequence 1, Appl
8	72	41.4	15	3	US-09-100-409A-38	Sequence 38, Appl
9	72	41.4	15	4	US-09-303-323-1	Sequence 1, Appl
10	72	41.4	15	5	PCT-US95-13841-12	Sequence 12, Appl
11	72	41.4	27	1	US-08-446-692-19	Sequence 19, Appl
12	72	41.4	27	2	US-08-488-351A-19	Sequence 19, Appl
13	72	41.4	27	3	US-09-100-414B-36	Sequence 36, Appl
14	72	41.4	27	4	US-09-303-323-36	Sequence 36, Appl
15	72	41.4	34	5	PCT-US95-13841-27	Sequence 27, Appl
16	72	41.4	35	1	US-08-446-692-55	Sequence 55, Appl
17	72	41.4	35	1	US-08-446-692-61	Sequence 61, Appl
18	72	41.4	35	2	US-08-488-351A-55	Sequence 55, Appl
19	72	41.4	35	2	US-08-488-351A-61	Sequence 61, Appl
20	72	41.4	36	4	US-09-082-279B-505	Sequence 505, Appl
21	72	41.4	36	4	US-09-315-304B-505	Sequence 505, Appl
22	72	41.4	42	5	PCT-US95-13841-25	Sequence 25, Appl
23	72	41.4	45	1	US-08-446-692-33	Sequence 33, Appl
24	72	41.4	45	2	US-08-488-351A-33	Sequence 33, Appl
25	72	41.4	438	3	US-08-486-099-105	Sequence 105, Appl
26	72	41.4	438	3	US-08-360-107A-115	Sequence 115, Appl
27	72	41.4	438	3	US-08-484-223B-105	Sequence 105, Appl

28 72 41.4 438 3 US-08-919-597-105 Sequence 105, App
29 72 41.4 438 3 US-08-475-668A-105 Sequence 105, App
30 72 41.4 438 3 US-08-485-551A-105 Sequence 105, App
31 72 41.4 438 3 US-08-471-913A-105 Sequence 105, App
32 72 41.4 438 4 US-08-485-264A-105 Sequence 105, App
33 72 41.4 438 4 US-08-474-349A-105 Sequence 105, App
34 72 41.4 550 1 US-08-279-700-16 Sequence 16, Appl
35 72 41.4 550 1 US-08-279-700-18 Sequence 18, Appl
36 72 41.4 550 1 US-08-279-700-20 Sequence 20, Appl
37 72 41.4 550 1 US-08-279-700-22 Sequence 22, Appl
38 72 41.4 550 1 US-08-348-891A-5 Sequence 5, Appl
39 72 41.4 550 2 US-08-905-817-5 Sequence 5, Appl
40 72 41.4 550 4 US-09-230-944-18 Sequence 18, Appl
41 72 41.4 550 4 US-09-230-944-20 Sequence 20, Appl
42 71 40.8 17 1 US-08-188-223-6 Sequence 6, Appl
43 71 40.8 17 4 US-08-968-466-6 Sequence 6, Appl
44 71 40.8 17 4 US-08-478-546B-6 Sequence 6, Appl
45 69 39.7 15 3 US-09-100-414B-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-460-502-10
; Sequence 10, Application US/08460502
; Patent No. 5843484
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Goltick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-502-10

Query Match 60.9%; Score 106; DB 2; Length 40;

Best Local Similarity 79.3%; Pred. No. 5.4e-09;

Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 6 GLRPGSGPSLKLSEIKGVIVHRLEGVE 34

DB 12 GTRDHHKGPGLKLSEIKGVIVHRLEGVE 40

```

1 RESULT 3
2 PCT-US95-13841-28
3 Sequence 28, Application PC/TUS9513841
4 GENERAL INFORMATION:
5 APPLICANT: United Biomedical Inc; Walfield, Alan M.;
6 APPLICANT: Wang, Chang Yi
7 TITLE OF INVENTION: Synthetic IgE Membrane Anchor
8 TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
9 NUMBER OF SEQUENCES: 32
10 CORRESPONDENCE ADDRESS:
11 ADDRESS: Maria C.H. Lin
12 STREET: 345 Park Avenue
13 CITY: New York
14 STATE: NY
15 COUNTRY: USA
16 ZIP: 10154
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: WordPerfect 5.1
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: PCT/US95/13841

```

APPELLANI, Wang, Chang et al.
TITLE OF INVENTION: Synthetic IgE Membrane Anchor
TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy

```
; MOLECULE TYPE: peptide
PCT-US95-13841-26

Query Match      42.2%; Score 73.5; DB 5; Length 43;
Best Local Similarity 64.3%; Pred. No. 0.00039;
Matches 18; Conservative 1; Mismatches 4; Indels 5; Gaps 1;

QY 6 GLRPGSSPSLKLSEIKGVIVHRLEGV 33
   | : | | | | | | | | | | | | | | | |
Db 21 GQQQGLGG-----LSEIKGVIVHRLEGV 43

RESULT 5
US-08-446-692-20
; Sequence 20, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-20

Query Match      42.0%; Score 73; DB 1; Length 42;
Best Local Similarity 93.8%; Pred. No. 0.00044;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 18 LSEIKGVIVHRLEGV 33
   : | | | | | | | | | | | | | | | |
Db 15 VLSEIKGVIVHRLEGV 30

RESULT 6
US-08-488-351A-20
; Sequence 20, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-20

Query Match      42.0%; Score 73; DB 2; Length 42;
Best Local Similarity 93.8%; Pred. No. 0.00044;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 18 LSEIKGVIVHRLEGV 33
   : | | | | | | | | | | | | | | | |
Db 15 VLSEIKGVIVHRLEGV 30

RESULT 7
US-09-100-414B-1
; Sequence 1, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
```

```

; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-1

```

```

Query Match 41.4%; Score 72; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 19 LSEIKGVIVHRLEGV 33
Db 1 LSEIKGVIVHRLEGV 15

```

```

RESULT 8
US-09-100-409A-38
; Sequence 38, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-409A-38

```

```

Query Match 41.4%; Score 72; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00018;

```

```

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 LSEIKGVIVHRLEGV 33
Db 1 LSEIKGVIVHRLEGV 15

```

```

RESULT 9
US-09-303-323-1
; Sequence 1, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-1

```

```

Query Match 41.4%; Score 72; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 19 LSEIKGVIVHRLEGV 33
Db 1 LSEIKGVIVHRLEGV 15

```

```

RESULT 10
PCT-US95-13841-12
; Sequence 12, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY

```


COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13841
FILING DATE: 25-OCT-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/328,519
FILING DATE: 25-OCT-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lin, Maria C.H.
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4117
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-13841-12

Query Match 41.4%; Score 72; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 11
US-08-446-692-19
Sequence 19, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELEPHONE: (212)415-8745

TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-19

Query Match 41.4%; Score 72; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 12
US-08-488-351A-19
Sequence 19, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-19

Query Match 41.4%; Score 72; DB 2; Length 27;

Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 13

US-09-100-414B-36
; Sequence 36, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-36

Query Match 41.4%; Score 72; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 14

US-09-303-323-36
; Sequence 36, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-36

Query Match 41.4%; Score 72; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 1 LSEIKGVIVHRLEGV 15

RESULT 15

PCT-US95-13841-27
; Sequence 27, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE/CHARACTERISTICS:

; LENGTH: 34 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

PCT-US95-13841-27

Query Match

41.4%; Score 72; DB 5; Length 34;

Best Local Similarity 100.0%; Pred. No. 0.00048;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33

Db 1 LSEIKGVIVHRLEGV 15

Search completed: October 10, 2002, 16:14:04

Job time : 8.19231 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 17.8718 Seconds
(without alignments)
211.311 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSGSLKLLSEIKGVIVHRLGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	60.9	40	AAW67581	Synthetic chimera f
2	106	60.9	40	AAW79986	Measles virus fusi
3	101	58.0	40	AAW67582	Synthetic chimera f
4	78	44.8	18	AAW35441	T-cell stimulatory
5	77	44.3	25	AAW62705	LHRH-containing im
6	74	42.5	20	AAW57161	Measles virus mmo
7	74	42.5	32	AAW57161	Measles virus F pr
8	73.5	42.2	74	AAW88391	migE2-GG-MVF1rh.
9	73.5	42.2	43	AAW05618	migE1-GG-MVF1rh.
10	73	42.0	42	AAW62708	LHRH-containing im
11	72	41.4	15	AAW62697	Helper T cell epit

12	72	41.4	15	AAW82591	Measles virus F pr
13	72	41.4	15	AAW82082	Measles virus glyc
14	72	41.4	15	AAW05604	Measles virus F pr
15	72	41.4	15	AAW88401	Measles virus F pr
16	72	41.4	15	AAW88392	Measles virus F pr
17	72	41.4	15	AAW80054	Pathogen derived T
18	72	41.4	15	AAW91121	Measles virus F pr
19	72	41.4	15	AAW44762	Measles virus prot
20	72	41.4	15	AAW68540	Helper T cell epit
21	72	41.4	15	AAW54537	T helper cell (Th)
22	72	41.4	15	AAW58764	Measles virus F pr
23	72	41.4	15	AAW84440	Amino acid sequenc
24	72	41.4	15	AAW86638	HER-2 B cell pepti
25	72	41.4	15	AAW62707	LHRH-containing im
26	72	41.4	27	AAW83584	IgE CH4 region con
27	72	41.4	27	AAW83576	IgE CH4 region con
28	72	41.4	27	AAW91156	MVF Th epitope/LHR
29	72	41.4	27	AAW68567	Peptide immunogen
30	72	41.4	34	AAW05619	MVF1rh-GG-migE2S.
31	72	41.4	35	AAW65381	Universal immunost
32	72	41.4	35	AAW65387	Universal immunost
33	72	41.4	36	AAW89143	Core polypeptide f
34	72	41.4	36	ABB00502	Viral Dp178/107-11
35	72	41.4	36	ABB01978	Viral core polypep
36	72	41.4	36	AAU13051	Dp178-like/Dp107-1
37	72	41.4	36	AAW77498	Core polypeptide T
38	72	41.4	42	AAW05617	MVF1rh-GG-migE1.
39	72	41.4	45	AAW62721	LHRH-containing im
40	72	41.4	436	AAU14096	Peptide sequence f
41	72	41.4	550	AAW34540	F protein of atten
42	72	41.4	550	AAW42394	Moraten haemagglut
43	72	41.4	550	AAW42395	San Diego haemaggl
44	72	41.4	550	AAW42396	Chicago 1 haemaggl
45	72	41.4	550	AAW42397	Consensus haemaggl

ALIGNMENTS

RESULT 1

AAW67581
ID AAW67581 standard; peptide; 40 AA.

XX
AC AAW67581;

XX
DT 02-MAR-1999 (first entry)

XX
DE Synthetic chimera fimbria/T-cell epitope peptide LBL.

XX
KW Chimeric; non-typable Haemophilus influenzae; fimbria; T-cell epitope;
KW immunogenic composition; immune response.

XX
OS Synthetic.

XX
PN US5843464-A.

XX
PD 01-DEC-1998.

XX
PF 02-JUN-1995; 95US-0460502.

XX
PR 02-JUN-1995; 95US-0460502.

XX
(OHIS) UNIV OHIO STATE.

XX
Bakaletz LO, Kaumaya PTP;

XX
WPI; 1999-044514/04.

XX
PT Synthetic chimeric fimbria peptide - useful for vaccination against
PT non-typable Haemophilus influenzae

XX
PS Claim 4; Column 4; 16pp; English.

KW T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
 KW scaffold; inhibition; metastasis; wound healing; solid phase.
 XX

OS Measles virus.
 XX

PN WO9738011-A1.
 XX

PP 16-OCT-1997.
 XX

PF 03-APR-1997; 97WO-DE00146.
 XX

PR 03-APR-1996; 96DK-0000398.
 XX

PA (PEPR-) PEPRESEARCH AS.
 XX

PI Heegaard PMH, Jakobsen PH;
 XX

DR WPI; 1997-512645/47.
 XX

PT Non-dendritic peptide carrier linked to a solid phase - useful as a
 PT diagnostic agent and as a scaffold for production of chemical
 PT derivatives
 XX

PS Claim 30; Page 199; 262pp; English.
 XX

XX A non-dendritic peptide carrier (A) has been developed which is coupled
 CC through a linker to a solid phase, forming a complex of (A)-solid phase.
 CC Where (A) comprises 10-50 amino acids capable of forming a secondary
 CC structure in a benign buffer after liberation from the solid phase, and
 CC further the (A)-solid phase complex comprises an immunogenic substance
 CC and/or an immune mediator coupled on (A). The present sequence
 CC represents a specifically claimed T-cell stimulatory peptide from the
 CC present invention. An (A)-solid phase complex can be used as a scaffold
 CC for the production of chemical derivatives, characterised by covalently
 CC attaching molecules at attachment points. Alternatively (A) is used as
 CC a scaffold-peptide for the incorporation into an immunostimulating
 CC complex (Iscom) resulting in an (A)-Iscom complex which is used for the
 CC chemical coupling of antigenic substances in an aqueous solution by
 CC conjugation. (A) derivatised with one or more peptides having
 CC fibronectin-, laminin- or vitronectin-like binding activities can be
 CC used for the promotion of cell-attachment to plastic surfaces, in
 CC particular to inhibit tumour growth and metastasis, and for promotion
 CC of wound healing. Also a derivatised (A) can be used for the selection
 CC of specifically-binding aptamers or as a diagnostic agent. Such
 CC diagnostic-(A) molecules could be used to detect molecules derived from
 CC or indicative of pregnancy or of a disease, such as an infectious,
 CC autoimmune or cancerous disease.
 XX

SQ Sequence 18 AA;

Query Match 44.8%; Score 78; DB 18; Length 18;
 Best Local Similarity 94.4%; Pred. No. 0.00014;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 KLLSEIKGVIVHRLEGVE 34
 ||||| ||||| ||||| |||||
 Db 1 KLLSLIKGVIVHRLEGVE 18

RESULT 5
 AAR62705
 ID AAR62705 standard; peptide; 25 AA.
 XX

AC AAR62705;
 XX

XX 10-SEP-1995 (first entry)
 XX

XX LHRH-containing immunogenic peptide.
 XX

KW Helper T cell epitope; universal immune stimulator; invasin; haptent;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW measles virus F protein.

XX Synthetic.
 XX

OS Key Location/Qualifiers
 XX 1..15

FT Domain /note= "measles virus F protein helper T cell epitope"
 XX 16..25

FT Domain /note= "LHRH haptent"
 XX

XX WO9425060-A.
 XX

XX 10-NOV-1994.
 XX

XX 28-APR-1994; 94WO-US04832.
 XX

XX 27-APR-1993; 93US-0057166.
 XX

XX 14-APR-1994; 94US-0229275.
 XX

XX (LADD/) LADD A E.
 XX (WANG/) WANG C Y.
 XX (ZAMB/) ZAMB T.
 XX

XX Ladd AE, Wang CY, Zamb T;
 XX

XX WPI; 1994-357910/44.
 XX

XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX

PS Claim 8; Page 84; 213pp; English.
 XX

XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptent containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and haptent
 CC components. When the haptent is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX

SQ Sequence 25 AA;

Query Match 44.3%; Score 77; DB 15; Length 25;
 Best Local Similarity 100.0%; Pred. No. 0.00029;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGVE 34
 ||||| ||||| ||||| |||||
 Db 1 LSEIKGVIVHRLEGVE 16

RESULT 6
 AAW57161
 ID AAW57161 standard; peptide; 20 AA.
 XX

XX AAW57161;
 XX

XX 30-JUL-1998 (first entry)
 XX

XX Measles virus mimotope for inhibiting binding of BH129 peptide 31.
 XX

KW Measles virus; haemagglutinin neutralising epitope region; MV isolate;

KW morbillivirus; vaccine; immune response; Edmonston strain; HNE;
 XX antigenic.

OS Synthetic.
 OS Measles virus.

PN WO9805682-A2.

XX 12-FEB-1998.

XX 26-JUL-1997; 97WO-EP04070.

XX 31-JUL-1996; 96EP-0112341.

XX (MULLER) MULLER C P.

XX Muller CP;

XX WPI; 1998-145549/13.

XX Morbillivirus antigenic determinant peptides - useful for, e.g.
 PT preparation of protective vaccines against measles

XX Claim 12; Fig 17; 67pp; English.

CC The present sequence represents a measles virus mimotope for inhibiting
 CC binding of BHL29 peptide. The peptide has an amino acid sequence
 CC corresponding to, or mimicking at least 1 antigenic determinant of the
 CC morbillivirus haemagglutinin protein. The peptide can be used to provoke
 CC an immune response to measles virus in mammals, and is useful in the
 CC preparation of a vaccine for the prevention of measles. The peptide is
 CC present in a vaccine at about 0.1 mu g to 100 mu g.

XX Sequence 20 AA;

Query Match 42.5%; Score 74; DB 19; Length 20;
 Best Local Similarity 88.9%; Pred. NO. 0.00062;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 16 LKLLSEIKGVIVHRLGV 33

DB 3 LSQLSEIKGVIVHRLGV 20

RESULT 7

AA888391
 ID AAR88391 standard; Protein; 72 AA.

XX AAR88391;

XX 12-JUN-1996 (first entry)

XX Measles virus F protein MVP-Cl.

XX Measles virus; myxo virus; F protein; T-antigen; vaccine; antibody;
 KW alpha-helix; coiled-coil heterodimer; core peptide; subunit;
 KW plasmid PHIL-S1/E.

OS Synthetic.

XX Key Location/Qualifiers
 FH Peptide 4..18

FT /label= mat.peptide
 FT /note= "measles virus F protein"

XX WO9531480-A1.

XX 23-NOV-1995.

XX 18-MAY-1995; 95WO-CA00293.

XX 18-MAY-1994; 94US-0245507.

XX

PA (SPIS-) SPI SYNTHETIC PEPTIDES INC.

XX Cachia PJ, Hodges RS, Houston ME, Irvin RT, Kay CM;
 PI Zhou NE;

XX WPI; 1996-010880/01.

DR N-PSDB; AAT09871.

XX Hetero:dimeric polypeptide immunogen in coiled-coil configuration
 PT with different antigens on each sub:unit - useful in vaccines and
 PT for antibody prodn.

XX Claim 4; Page 71; 95pp; English.

XX This protein, encoded by DNA cloned in the polylinker region of plasmid
 CC PHIL-S1/E, encodes the measles virus F protein MVP-C1 which may be
 CC used to form one of the single chain polypeptide subunits of a
 CC synthetic vaccine. This antigen may be expressed in tandem with a
 CC carrier polypeptide, such that they form a single polypeptide chain,
 CC to form a decorated peptide.

XX Sequence 72 AA;

Query Match 42.5%; Score 74; DB 17; Length 72;

Best Local Similarity 83.3%; Pred. NO. 0.0029;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 17 KLLSEIKGVIVHRLGVE 34

DB 2 EFLSEIKGVIVHRLGVK 19

RESULT 8

AAW05620

ID AAW05620 standard; peptide; 34 AA.

XX AAW05620;

XX 10-DEC-1996 (first entry)

XX mIgE2-GG-MVF1Th.

XX Immunoglobulin; IgE; membrane protein; human; epsilon chain; hepatitis B;
 KW membrane anchoring domain; helper T cell; surface antigen; core antigen;
 KW pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;
 KW Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;
 KW diphtheria toxin; plasmodium falciparum; circumsporozoite; E. coli Trat;
 KW schistosoma mansoni; triose phosphate isomerase; allergic reaction;
 KW allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;
 KW antihistamine; decongestant; beta-2 agonist; immunosuppression;
 KW corticosteroid.

XX Synthetic.

XX Key Location/Qualifiers
 FH Peptide 1..17

FT /note= "immunoglobulin E fragment"

FT Peptide 20..34

XX /note= "measles virus F protein helper T cell epitope #1"

XX WO9612740-A1.

XX 02-MAY-1996.

XX 25-OCT-1995; 95WO-US13841.

XX 25-OCT-1994; 94US-0328519.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Walfield AM, Wang CY;

XX WPI; 1996-230555/23.

PR 27-APR-1993; 93US-0057166.
 PR 14-APR-1994; 94US-0229275.
 XX (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 PI WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 DR that suppress LHRH activity in males and females
 XX Claim 8; Page 86; 213pp; English.
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasin-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX SQ Sequence 42 AA;
 Query Match 42.0%; Score 73; DB 15; Length 42;
 Best Local Similarity 93.8%; Pred. No. 0.0021;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 18 LLEIKGVIVHRLEGV 33
 :|||||
 DB 15 VLSEIKGVIVHRLEGV 30
 RESULT 11
 AAR62697
 ID AAR62697 standard; peptide; 15 AA.
 XX AAR62697;
 AC 10-SEP-1995 (first entry)
 DT Helper T cell epitope for use in universal immune stimulator.
 DE Helper T cell epitope; universal immune stimulator; invasin; hapten;
 KW vaccine; measles.
 KW Measles virus.
 OS WO9425060-A.
 PN 10-NOV-1994.
 PD 28-APR-1994; 94WO-US04832.
 XX 27-APR-1993; 93US-0057166.
 PR 14-APR-1994; 94US-0229275.
 XX (LADD/) LADD A E.
 PA (WANG/) WANG C Y.

PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 DR Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 PT Claim 7; Page 25; 213pp; English.
 PS Synthetic immunogenic peptides are provided in which a universal immune
 XX stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents a measles virus F protein helper T cell
 CC epitope which can be used as Th in the immune stimulator.
 XX SQ Sequence 15 AA;
 Query Match 41.4%; Score 72; DB 15; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 19 LSEIKGVIVHRLEGV 33
 :|||||
 DB 1 LSEIKGVIVHRLEGV 15
 RESULT 12
 AAR82591
 ID AAR82591 standard; peptide; 15 AA.
 XX AAR82591;
 AC 13-JUN-1996 (first entry)
 DT Measles virus F protein cell epitope, MVF1.
 DE IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 KW vaccine; allergy; antibody; constant heavy chain.
 KW Measles virus.
 OS WO9526365-A1.
 PN 05-OCT-1995.
 PD 24-MAR-1995; 95WO-US03741.
 XX 25-OCT-1994; 94US-0328912.
 PR 28-MAR-1994; 94US-0218461.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA Wang CY;
 XX WPI; 1995-351297/45.
 DR Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper
 PT T cell epitope - useful for eliciting antibody prodn. for allergy
 PT treatment
 XX Claim 3; Page 22; 87pp; English.

CC AAR82571₉₁ are helper T cell epitopes which can be used in the
 CC preparation of a peptide immunogen that is useful in vaccines for
 CC treating allergic reactions. In the immunogen an IgE CH4 peptide
 CC is attached C-terminally to a series of amino acids including a
 CC helper T cell epitope. The immunogen may also opt. contain a fatty
 CC acid or fatty acid derivative, an invasin domain or alpha-NH2. The
 CC immunogen produces high titres of antibodies to the effector site
 CC in human IgE heavy chain (the CH4 domain peptide) which inhibit mast
 CC cell activation and reduce allergen-induced IgE prodn. The immunogens
 CC may be used in either a radially branching multimeric form or a
 CC linearly arranged monomeric form.

XX Sequence 15 AA;

Query Match 41.4%; Score 72; DB 16; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33

Db 1 LSEIKGVIVHRLEGV 15

RESULT 13

AAR82082

ID AAR82082 standard; peptide; 15 AA.

XX AC AAR82082;

DT 03-APR-1996 (first entry)

DE Measles virus glycoprotein F 288-302 T cell epitope.

XX Retro-inverso modified; T cell epitope; analogue; vaccine; B cell;
 KW immunisation; MVF; glycoprotein F; gpF; 288-302.

OS Measles virus.

PN WO9523166-A1.

XX 31-AUG-1995.

XX 24-FEB-1995; 95WO-AU00090.

XX 25-FEB-1994; 94AU-0004119.

XX (DEAK-) DEAKIN RES LTD.

XX Comis A, Fischer P, Tyler MT;

XX WPI; 1995-311503/40.

XX Synthetic peptide T cell epitope analogue (retro-)inverso modified
 PT - used conjugated with a B cell epitope in a vaccine tailored to a
 PT specific condition, e.g. polio, hepatitis B, etc.

XX Claim 3; Page 38; 51pp; English.

XX A vaccine comprising a retro-inverso modified, native T cell epitope
 CC i.e. AAR82073-R82094, conjugated to a B-cell epitope (i.e. the malaria
 CC immunodominant B-cell epitope (AAR82095), and epitopes based on
 CC poliomyelitis, tetanus or hepatitis B proteins etc.) is tailored
 CC to a condition of interest (depending on the epitopes used), and
 CC used to immunise a host against the specific condition.

XX Sequence 15 AA;

Query Match 41.4%; Score 72; DB 16; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33

Db 1 LSEIKGVIVHRLEGV 15

Db 1 LSEIKGVIVHRLEGV 15

RESULT 14

AAW05604

ID AAW05604 standard; peptide; 15 AA.

XX AC AAW05604;

XX 10-DEC-1996 (first entry)

DE Measles virus F protein helper T cell epitope #1.

XX Immunoglobulin; IgE; membrane protein; human; epsilon chain; hepatitis B;
 KW membrane anchoring domain; helper T cell; surface antigen; core antigen;
 KW pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;
 KW Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;
 KW diphtheria toxin; plasmodium falciparum; circumsporozoite; E. coli Trat;
 KW schistosoma mansoni; triose phosphate isomerase; allergenic reaction;
 KW allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;
 KW antihistamine; decongestant; beta-2 agonist; immunosuppression;
 KW corticosteroid.

XX Measles virus.

XX WO9612740-A1.

XX 02-MAY-1996.

XX 25-OCT-1995; 95WO-US13841.

XX 25-OCT-1994; 94US-0328519.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Walfield AM, Wang CY;

XX WPI; 1996-230555/23.

XX Peptide immunogen useful in treatment of allergy - comprises
 PT membrane-bound IgE epsilon-chain peptide synthesised linearly in
 PT tandem with T helper epitope peptide

XX Claim 2; Page 18; 53pp; English.

XX AAW05957-W05616 represent helper T cell epitopes used in the peptide
 CC immunogens of the invention. This sequence represents a measles virus F
 CC protein helper T cell antigen. The peptides of the invention contain
 CC one of these sequences, and a membrane-bound immunoglobulin E (IgE)
 CC fragment (see AAW05595 and AAW05596). The peptide immunogens of the
 CC invention can be used in vaccines for the immunotherapeutic treatment of
 CC allergic reactions, including allergic rhinitis, food allergies,
 CC anaphylaxis, or virally-induced asthma. The immunogens overcome the
 CC short effective period of antihistamines, decongestants, and beta-2
 CC agonists, while preventing the broad immunosuppression of
 CC corticosteroids. The peptides do not have the potential side effects of
 CC restlessness or sedation (associated with antihistamines), associated
 CC increased morbidity in asthmatics (as seen with beta-2 agonists) and
 CC adverse hormonal activities (observed in corticosteroid users).

XX Sequence 15 AA;

Query Match 41.4%; Score 72; DB 17; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33

Db 1 LSEIKGVIVHRLEGV 15

RESULT 15

AAR88401

ID AAR88401 standard; Peptide; 15 AA.
 AC AAR88401;
 XX
 XX
 DT 22-JUN-1996 (first entry)
 XX
 DE Measles virus F protein MVF peptide B-cell antigen.
 XX
 KW B-antigen; vaccine; antibody; T-cell; T-lymphocyte;
 alpha-helix; coiled-coil heterodimer; core peptide; subunit.
 XX
 OS Synthetic.
 XX
 PN W09531480-A1.
 XX
 PD 23-NOV-1995.
 XX
 PF 18-MAY-1995; 95WO-CA00293.
 XX
 PR 18-MAY-1994; 94US-0245507.
 XX
 PA (SPIS-) SPI SYNTHETIC PEPTIDES INC.
 XX
 PI Cachia PJ, Hodges RS, Houston ME, Irvin RT, Kay CM;
 PI Zhou NE;
 XX
 DR WPI; 1996-010880/01.
 XX
 PT Hetero:dimeric polypeptide immunogen in coiled-coil configuration
 PT with different antigens on each sub:unit - useful in vaccines and
 PT for antibody prodn.
 XX
 PS Disclosure; Page 63; 95pp; English.
 XX
 CC This antigenic peptide may be attached to a core peptide
 CC contained in one of the 2 subunits of an alpha-helical coiled-coil
 CC heterodimer. The other subunit may contain e.g. a T-cell antigen
 CC peptide. Each core peptide is comprised of terminal and
 CC internal AA repeat sequences. This peptide antigen is attached
 CC to the core peptide through covalent linkages to certain AA of the
 CC internal repeats. The 2 subunits of the heterodimer are arranged
 CC in a stable alpha-helical coiled-coil configuration having a 1:1
 CC stoichiometry, and the peptide antigen is disposed toward the outer
 CC surfaces of the configuration. The heterodimer may be used as a
 CC synthetic vaccine (optionally multivalent) or to generate
 CC antibodies.
 XX
 SQ Sequence 15 AA;
 Query Match 41.4%; Score 72; DB 17; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00085;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 LSEIKGVIVHRLEGV 33
 Db 1 LSEIKGVIVHRLEGV 15

Search completed: October 10, 2002, 16:05:12
 Job time : 18.8718 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 7.9188 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHWSYGLRPGSSPSLKLLSEIKGVIVHRLEGVE 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR71: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	72	41.4	282	2 PQ0376	cell fusion glycop
2	72	41.4	282	2 PQ0388	cell fusion glycop
3	72	41.4	534	1 JU0274	cell fusion glycop
4	72	41.4	546	2 S47300	gene F protein - r
5	72	41.4	550	1 E48556	cell fusion glycop
6	72	41.4	553	1 VGNZMV	cell fusion glycop
7	71	40.8	546	1 VGNZRK	cell fusion glycop
8	71	40.8	546	2 S47305	gene F protein - r
9	66	37.9	542	2 JQ2223	cell fusion protei
10	66	37.9	546	1 VGNZRL	cell fusion glycop
11	66	37.9	662	1 VGNZCD	cell fusion glycop
12	66	37.9	662	2 S21382	cell fusion protei
13	65	37.4	552	2 S47034	cell fusion protei
14	65	37.4	631	1 VGNZPD	cell fusion glycop
15	65	37.4	631	1 A48346	cell fusion glycop
16	64	36.8	67	2 I78541	gonadoliberin prec
17	64	36.8	92	1 RHUG	gonadoliberin prec
18	62	35.6	89	2 I51423	gonadoliberin prec
19	60	34.5	546	2 S53386	cell fusion protei
20	58	33.3	10	1 RHFG	gonadoliberin - pi
21	58	33.3	10	1 RHSHG	gonadoliberin - sh
22	58	33.3	90	1 RHMSG	gonadoliberin prec
23	58	33.3	92	1 RHRTG	gonadoliberin prec
24	56.5	32.5	98	2 I50739	gonadotropin-relea
25	56	32.2	636	2 S47299	gene F protein - r
26	55	31.6	92	2 I50644	gonadoliberin I pr
27	54.5	31.3	80	2 S39779	aldehyde reductase
28	54.5	31.3	249	2 A41497	36K antigen pra -
29	54	31.0	10	1 RHAQ1	gonadoliberin I -

30 52 29.9 213 1 EPBO
31 51 29.3 75 2 AI3191
32 51 29.3 201 2 T47113
33 51 29.3 338 2 H87537
34 51 29.3 371 2 T49908
35 50.5 29.0 240 2 D70894
36 50 28.7 44 2 S28770
37 50 28.7 133 2 H86379
38 50 28.7 317 2 H97192
39 50 28.7 392 2 A69474
40 50 28.7 719 2 T52510
41 50 28.7 1132 2 AD1809
42 49.5 28.4 342 1 QOECB8
43 49 28.2 286 2 E82202
44 49 28.2 315 2 A88043
45 49 28.2 316 2 C84635

ALIGNMENTS

RESULT 1

PQ0376
cell fusion glycoprotein - measles virus (strain TT) (fragment)
C:Species: measles virus
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C:Accession: PQ0376
R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compari
A:Reference number: PQ0374; MUID:92300360
A:Accession: PQ0376
A:Molecule type: genomic RNA
A:Residues: 1-282 <SCH>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion

Query Match 41.4%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
DB 20 LSEIKGVIVHRLEGV 34

RESULT 2

PQ0388
cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)
C:Species: measles virus
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999
C:Accession: PQ0388
R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Weiss, R.A.
J. Gen. Virol. 73, 1581-1586, 1992
A:Title: A measles virus isolate from a child with Kawasaki disease: sequence compari
A:Reference number: PQ0374; MUID:92300360
A:Accession: PQ0388
A:Molecule type: genomic RNA
A:Residues: 1-282 <SCH>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion

Query Match 41.4%; Score 72; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
DB 20 LSEIKGVIVHRLEGV 34

RESULT 3

JU0274
 cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain Y
 N:Contains: fusion glycoprotein F1; fusion glycoprotein F2
 C:Species: subacute sclerosing panencephalitis virus, SSPEV
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jun-2000
 C:Accession: JU0274
 R:Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.
 Virus Genes 4, 173-181, 1990
 A:Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe
 A:Reference number: JU0274; MUID:90385702
 A:Accession: JU0274
 A:Molecule type: mRNA
 A:Residues: 1-534 <KOM>
 A:Cross-references: EMBL:D10548; NID:9222256; PIDN:BAA01405.1; PID:9222257
 A:Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue
 C:Genetics:
 A:Gene: F
 C:Superfamily: parainfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein
 F:1-22/Domain: signal sequence #status predicted <Sig>
 F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
 F:108-534/Product: cell fusion glycoprotein F1 #status predicted <FF1>
 F:498-514/Domain: transmembrane #status predicted <TMN>
 F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.4%; Score 72; DB 1; Length 534;

Best Local Similarity 100.0%; Pred. No. 0.046;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

DB 288 LSEIKGVIVHRLGV 302

RESULT 4

S47300
 gene F protein - rinderpest virus
 C:Species: rinderpest virus
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
 C:Accession: S47300; PQ0865
 R:Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.
 submitted to the EMBL Data Library, March 1994
 A:Description: The complete nucleotide sequence of the fusion protein gene of the vaccin
 A:Reference number: S47299
 A:Accession: S47300
 A:Molecule type: DNA
 A:Residues: 1-546 <EVA>
 A:Cross-references: EMBL:Z31656; NID:9535406; PIDN:CAA83482.1; PID:9535407
 R:Chamberlain, R.W.; Wamwayi, H.M.; Hockley, E.; Shalla, M.S.; Goatley, L.; Knowles, N.J
 J. Gen. Virol. 74, 2775-2780, 1993
 A:Title: Evidence for different lineages of rinderpest virus reflecting their geographic
 A:Reference number: PQ0865; MUID:94103786
 A:Accession: PQ0865
 A:Molecule type: mRNA
 A:Residues: 86-191 <CHA>
 C:Genetics:
 A:Gene: F
 C:Superfamily: parainfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 41.4%; Score 72; DB 2; Length 546;

Best Local Similarity 100.0%; Pred. No. 0.047;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

DB 284 LSEIKGVIVHRLGV 298

RESULT 5

C:Superfamily: parainfluenza virus cell fusion protein

E48556

cell fusion glycoprotein precursor - measles virus (strain AIK-C)
 C:Species: measles virus
 C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
 C:Accession: E48556
 R:Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.
 Virus Genes 7, 67-81, 1993
 A:Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the A-K
 A:Reference number: A48556; MUID:93227570
 A:Accession: E48556
 A:Molecule type: genomic RNA
 A:Residues: 1-550 <MOR>
 A:Cross-references: GB:S58435; NID:9299460; PIDN:AAB26145.1; PID:9299465
 A:Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBI:P:129272)
 C:Genetics:
 A:Gene: F
 C:Superfamily: parainfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein
 F:1-22/Domain: signal sequence #status predicted <Sig>
 F:23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>
 F:108-550/Product: cell fusion glycoprotein F1 #status predicted <FF1>
 F:113-138/Region: hydrophobic
 F:495-514/Domain: transmembrane #status predicted <TMN>
 F:6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.4%; Score 72; DB 1; Length 550;

Best Local Similarity 100.0%; Pred. No. 0.048;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33

DB 288 LSEIKGVIVHRLGV 302

RESULT 6

VCN2MV

cell fusion glycoprotein precursor - measles virus

C:Species: measles virus
 C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000
 C:Accession: A26962; A25616; PQ0380; PQ0384
 R:Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.
 J. Gen. Virol. 68, 1695-1703, 1987
 A:Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and
 A:Reference number: A92794; MUID:87224816
 A:Accession: A26962
 A:Molecule type: mRNA
 A:Residues: 1-553 <BUC>
 A:Cross-references: GB:D00090; NID:g222061; PIDN:BAA00056.1; PID:g222062
 A:Experimental source: strain Halle
 R:Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Belli
 Virology 155, 508-523, 1986
 A:Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles
 A:Reference number: A94350; MUID:87071668
 A:Accession: A25616
 A:Molecule type: mRNA
 A:Residues: 4-553 <RIC>
 A:Cross-references: GB:M14915; NID:g331762; PIDN:AAA46423.1; PID:g331763
 A:Experimental source: strain Edmonston
 R:Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.
 J. Gen. Virol. 73, 1581-1586, 1992
 A:Title: A measles virus isolate from a child with Kawasaki disease: sequence comp
 A:Reference number: PQ0374; MUID:92300360
 A:Accession: PQ0380
 A:Molecule type: genomic RNA
 A:Residues: 272-553 <SCH1>
 A:Experimental source: isolate CL
 A:Accession: PQ0384
 A:Molecule type: genomic RNA
 A:Residues: 272-553 <SCH2>
 A:Experimental source: isolate SE
 C:Genetics:
 A:Gene: F
 C:Superfamily: parainfluenza virus cell fusion protein

C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-25/Domain; signal sequence #status predicted <SIG>
F:26-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>
F:111-553/Product: cell fusion glycoprotein F1 #status predicted <FF1>
F:501-517/Domain: transmembrane #status predicted <TMN>
F:32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.4%; Score 72; DB 1; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLGV 33
|||||:|||||

Db 291 LSEIKGVIVHRLGV 305

RESULT 7
VGNZRL
cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: rinderpest virus
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 25-Oct-1996
C;Accession: A31051
R;Hsu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.
Virology 166, 149-153, 1988
A;Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis
A;Reference number: A31051; MUID:88322864
A;Accession: A31051
A;Molecule type: genomic RNA
A;Residues: 1-546 <HSD>
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-108/Product: cell fusion glycoprotein F2 #status predicted <FF1>
F:109-546/Product: cell fusion glycoprotein F1 #status predicted <FF2>
F:109-134/Domain: transmembrane #status predicted <TN1>
F:491-513/Domain: transmembrane #status predicted <TN2>
F:25,57,63,518/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.8%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.065;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLGV 33
|||||:|||||

Db 284 LSEIKGVIVHRLGV 298

RESULT 8
S47305
gene F protein - rinderpest virus
C;Species: rinderpest virus
C;Date: 20-Oct-1994 #sequence_revision 08-Sep-1995 #text_change 20-Sep-1999
C;Accession: S47305; S47301
R;Baron, M.D.; Barrett, T.
submitted to the EMBL Data Library, March 1994
A;Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30
A;Reference number: S47283
A;Accession: S47305
A;Molecule type: mRNA
A;Residues: 1-546 <BAR>
A;Cross-references: EMBL:230697; NID:9535396; PIDN:CAAB3181.1; PID:9535401; EMBL:230700;
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: transmembrane protein

Query Match 40.8%; Score 71; DB 2; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.065;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLGV 33
|||||:|||||

Db 284 LSEIKGVIVHRLGV 298

RESULT 9
JQ2223
cell fusion protein F0 precursor - phocine distemper virus
N;Contains: F1 and F2 chains
C;Species: phocine distemper virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 24-Nov-1999
C;Accession: JQ2223
R;Visser, I.K.G.; van der Heijden, R.W.J.; van de Bilt, M.W.G.; Kenter, M.J.H.; Oerv
J. Gen. Virol. 74, 1989-1994, 1993
A;Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites
e virus entity.
A;Reference number: JQ2223; MUID:93389459
A;Accession: JQ2223
A;Molecule type: mRNA
A;Residues: 1-542 <VIS>
A;Cross-references: GB:L07075
A;Note: The authors translated the codon ATC for residue 4 as Leu
C;Comment: This fusion protein F0 is cleaved into F1 and F2 chains.
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-542/Product: fusion protein #status predicted <MAT>
F:16-99/Product: F2 chain #status predicted <F2C>
F:105-542/Product: F1 chain #status predicted <F1C>
F:105-135/Region: hydrophobic
F:486-512/Domain: transmembrane #status predicted <TMN>
F:21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.9%; Score 66; DB 2; Length 542;
Best Local Similarity 73.7%; Pred. No. 0.33;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 15 SUKLLSEIKGVIVHRLGV 33
|:|||||:|||||

Db 276 SYPTLSEIKGVIVHRLGV 294

RESULT 10
VGNZRL
cell fusion glycoprotein precursor - rinderpest virus (strain L)
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2
C;Species: rinderpest virus
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C;Accession: A28921
R;Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.
Virology 164, 523-530, 1988
A;Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of t
A;Reference number: A28921; MUID:88219541
A;Accession: A28921
A;Molecule type: mRNA
A;Residues: 1-546 <TSU>
A;Cross-references: GB:M20870; NID:g333898; PIDN:AAA47399.1; PID:g333899
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>
F:105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>
F:109-133/Domain: transmembrane #status predicted <TN1>
F:485-513/Domain: transmembrane #status predicted <TN2>
F:25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.9%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.33;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLGV 33

```
Db 284 LSEIKGVIVHRLEGV 298
|||||
RESULT 11
VGNZCD
cell fusion glycoprotein precursor - canine distemper virus
N:Contains: fusion protein F1; fusion protein F2
C:Species: canine distemper virus
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: JS0321
R:Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.
Virus Res. 8, 373-386, 1987
A:Title: The nucleotide sequence of the gene encoding the F protein of canine distemper
A:Reference number: JS0321; MUID:88129050
A:Accession: JS0321
A:Molecule type: mRNA
A:Residues: 1-662 <BAR>
A:Cross-references: GB:M21849; NID:g323241; PIDN:AAA42878.1; PID:g323242
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-135/Domain: signal sequence #status predicted <SIG>
F:136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>
F:225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>
F:606-629/Domain: transmembrane #status predicted <MEM>
F:62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 37.9%; Score 66; DB 1; Length 662;
Best Local Similarity 73.7%; Pred. No. 0.41;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 15 SLKLLSEIKGVIVHRLEGV 33
| |||:|||||
Db 396 SYPTLSEVKGIVVHRLEAV 414
|||||
RESULT 12
S21382
cell fusion protein - canine distemper virus
C:Species: canine distemper virus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S21382
R:Wild, T.F.; Bernard, A.; Spehner, D.; Villevall, D.; Drillien, R.
submitted to the EMBL Data Library, April 1992
A:Description: Vaccination of mice against canine distemper virus induced encephalitis
A:Reference number: S21382
A:Accession: S21382
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-662 <WIL>
A:Cross-references: EMBL:X65509; NID:g58853; PIDN:CAA46481.1; PID:g58854
C:Superfamily: parainfluenza virus cell fusion protein
Query Match 37.9%; Score 66; DB 2; Length 662;
Best Local Similarity 73.7%; Pred. No. 0.41;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 15 SLKLLSEIKGVIVHRLEGV 33
| |||:|||||
Db 396 SYPTLSEVKGIVVHRLEAV 414
|||||
RESULT 13
S47034
cell fusion protein precursor - porpoise morbillivirus
N:Alternate names: F protein
C:Species: porpoise morbillivirus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S47034
R:Bolt, G.; Gortschack, E.; Blixenkron-Moeller, M.; Wishaupt, R.G.A.; Welsh, M.J.; Ea
submitted to the EMBL Data Library, July 1994
A:Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbilli
A:Reference number: S47034
A:Accession: S47034
A:Molecule type: mRNA
A:Residues: 1-552 <BOL>
A:Cross-references: EMBL:X80757; NID:g520639; PIDN:CAA56731.1; PID:g520640
A:Experimental source: isolate Ulster 88
A:Note: the source is isolated as Cetacean morbillivirus
C:Superfamily: parainfluenza virus cell fusion protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-552/Product: fusion protein #status predicted <MAT>
Query Match 37.4%; Score 65; DB 2; Length 552;
Best Local Similarity 86.7%; Pred. No. 0.46;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 19 LSEIKGVIVHRLEGV 33
| |||:|||||
Db 290 LSEVKGIVVHRLEAV 304
|||||
RESULT 14
VGNZPD
cell fusion glycoprotein precursor - phocine distemper virus
N:Contains: fusion protein F1; fusion protein F2
C:Species: phocine distemper virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 25-Oct-1996
C:Accession: JQ1368
R:Koevaanes, J.; Blixenkron-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.
J. Gen. Virol. 72, 2959-2966, 1991
A:Title: The nucleotide sequence and deduced amino acid composition of the haemagglit
A:Reference number: JQ1368; MUID:92113538
A:Accession: JQ1368
A:Molecule type: genomic RNA
A:Residues: 1-631 <KOV>
C:Genetics:
A:Gene: F
C:Superfamily: parainfluenza virus cell fusion protein
C:Keywords: glycoprotein; membrane fusion; transmembrane protein
F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F:89-106/Domain: transmembrane #status predicted <TM1>
F:189-193/Region: cleavage processing #status predicted
F:194-631/Product: cell fusion glycoprotein F1 #status predicted <F1P>
F:194-212/Domain: transmembrane #status predicted <TM2>
F:575-595/Domain: transmembrane #status predicted <TM3>
F:110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 37.4%; Score 65; DB 1; Length 631;
Best Local Similarity 68.4%; Pred. No. 0.54;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 15 SLKLLSEIKGVIVHRLEGV 33
| |||:|||||
Db 365 SYPTLSEVKGIVVHRLEAV 383
|||||
RESULT 15
A48346
cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)
N:Contains: fusion protein F1; fusion protein F2
C:Species: phocine distemper virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 25-Oct-1996
C:Accession: A48346
R:Curran, M.D.; Lu, Y.J.; Rima, B.K.
Arch. Virol. 126, 159-169, 1992
A:Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced a
A:Reference number: A48346; MUID:92398437
A:Accession: A48346
A:Molecule type: mRNA
A:Residues: 1-631 <CUR>
A:Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBIP:113099)
C:Genetics:
A:Gene: F
```


C:Superfamily::parainfluenza virus cell fusion protein
 C:Keywords: glycoprotein; membrane fusion; transmembrane protein
 F:1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
 F:89-106/Domain: transmembrane #status predicted <TM1>
 F:194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
 F:194-219/Domain: transmembrane #status predicted <TM2>
 F:575-595/Domain: transmembrane #status predicted <TM3>
 F:110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 37.4%; Score 65; DB 1; Length 631;
 Best Local Similarity 68.4%; Pred. NO. 0.54;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 SLKLLSEIKGVIVHRLGV 33
 DB 365 SYPTLSEVKGVVVHRLGV 383

Search completed: October 10, 2002, 16:12:09
 Job time : 9.99573 secs

* GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 4.43162 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-13

Perfect score: 174

Sequence: 1 XHWSYGLRPSSGSLKLLSEIKGVIVHRLGVE 34

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	72	41.4	534	1	VGLF_MEASY
2	72	41.4	546	1	VGLF_RINDB
3	72	41.4	550	1	VGLF_MEASA
4	72	41.4	550	1	VGLF_MEASE
5	71	40.8	546	1	VGLF_RINDR
6	66	37.9	546	1	VGLF_RINDL
7	66	37.9	662	1	VGLF_CDVO
8	65.5	37.6	95	1	GON1_SPAAU
9	65	37.4	546	1	VGLF_RINDX
10	65	37.4	631	1	VGLF_PHODV
11	64	36.8	67	1	GON1_MACMU
12	64	36.8	92	1	GON1_HUMAN
13	64	36.8	529	1	VGLF_MEASI
14	62.5	35.9	95	1	GON1_PAGNA
15	62	35.6	89	1	GON1_XENLA
16	61.5	35.3	61	1	GON1_SHEEP
17	60	34.5	92	1	GON1_TUPGB
18	58	33.3	63	1	GON1_MEASU
19	58	33.3	90	1	GON1_MOUSE
20	58	33.3	91	1	GON1_PIG
21	58	33.3	92	1	GON1_RAT
22	58	33.3	99	1	GON1_DICLA
23	57	32.8	95	1	GON1_MORSA
24	56.5	32.5	94	1	GON1_HAPBU
25	55	31.6	92	1	GON1_CHICK
26	54.5	31.3	249	1	PRA_MYCLE
27	54	31.0	10	1	GON1_ALLMT
28	52	29.9	213	1	PBCT_BOVIN
29	51.5	29.6	74	1	GON3_ONCMY
30	51.5	29.6	90	1	GON8_RANDY
31	51	29.3	393	1	FEZ1_RAT
32	50.5	29.0	240	1	PRA_MVCTU
33	50	28.7	1131	1	APCE_ANASP

34	50	28.7	2114	1	MY9B_MOUSE
35	49.5	28.4	342	1	MURB_ECOLI
36	49	28.2	92	1	GON1_CAVPO
37	49	28.2	521	1	UBP3_HUMAN
38	48.5	27.9	359	1	PTN7_RAT
39	48	27.6	80	1	GON1_CLAGA
40	48	27.6	336	1	UL34_EBV
41	48	27.6	360	1	PTN7_HUMAN
42	48	27.6	395	1	PTP_ENCCU
43	47.5	27.3	74	1	GON3_ONCTS
44	47.5	27.3	82	1	GON3_SALSA
45	47.5	27.3	82	1	GON3_SALTR

ALIGNMENTS

RESULT 1					
VGLF_MEASY					
ID	VGLF_MEASY	STANDARD;	PRT;	534	AA.
AC	P26032;				
DT	01-MAY-1992 (Rel. 22, Created)				
DT	01-MAY-1992 (Rel. 22, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;				
DE	Fusion glycoprotein F1].				
GN	F.				
OS	Measles virus (strain Yamagata-1) (Subacute sclerose panencephalitis virus).				
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;				
OC	Paramyxoviridae; Paramyxovirinae; Morbillivirus.				
OX	NCBI_TaxID=11239;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90385702; Pubmed=1698327;				
RA	Komase K., Haga T., Yoshikawa Y., Sato T.A., Yamanouchi K.;				
RT	"Molecular analysis of structural protein genes of the Yamagata-1 strain of defective subacute sclerosing panencephalitis virus. IV. Nucleotide sequence of the fusion gene."				
RT	Virus Genes 4:173-181(1990).				
RL	-!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR MEMBRANES.				
CC	-!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2 LINKED BY A DISULFIDE BOND.				
CC	-!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; D10548; BAA01405.1; -				
DR	PIR; J00274; J00274.				
DR	HSSP; P04849; ISVF.				
DR	InterPro; IPR000776; Fusion_gly.				
DR	Pfam; PF00523; fusion_gly; 1.				
KW	Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.				
FT	SIGNAL 1 23				
FT	CHAIN 24 534				
FT	CHAIN 24 112				
FT	CHAIN 113 534				
FT	TRANSMEM 113 136				
FT	DOMAIN 137 494				
FT	TRANSMEM 495 515				
FT	DOMAIN 516 534				
FT	DISULFID 68 195				
FT	CARBOHYD 29 29				
FT	CARBOHYD 61 61				
FT	CARBOHYD 67 67				

```
SQ SEQUENCE 534 AA; 57963 MW; F5B21757E643844D CRC64;
Query Match 41.4%; Score 72; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 2
VGLF_RINDB STANDARD; PRT; 546 AA.
AC P41360; 1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
Fusion glycoprotein F1].
GN F.
OS Measles virus (strain AIK-C) (Subacute sclerosing panencephalitis
virus)
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=39007;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9508609; PubMed=7996154;
RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus.";
RL J. Gen. Virol. 75:3611-3617(1994).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z31656; CRA83482.1; -.
CC DR HSP; P04849; 1SVF.
CC DR InterPro; IPR000776; Fusion_gly.
CC Pfam; PF00523; fusion_gly; 1.
CC KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
CC SIGNAL 1 19
CC FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
CC FT CHAIN 20 108 F2 PROTEIN.
CC FT CHAIN 109 546 F1 PROTEIN.
CC FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
CC FT TRANSMEM 109 133 POTENTIAL.
CC FT TRANSMEM 484 513 POTENTIAL.
CC FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
CC FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
CC FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 518 518 O-LINKED (POTENTIAL).
CC SEQUENCE 546 AA; 58418 MW; 38B539B89344F401 CRC64;

Query Match 41.4%; Score 72; DB 1; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.0082;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 4
VGLF_MEASE STANDARD; PRT; 550 AA.
AC P08300;

Db 284 LSEIKGVIVHRLEGV 298

RESULT 3
VGLF_MEASA STANDARD; PRT; 550 AA.
AC P35973;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
Fusion glycoprotein F1].
GN F.
OS Measles virus (strain AIK-C) (Subacute sclerosing panencephalitis
virus)
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=36408;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227570; PubMed=8470368;
RA Mori T., Sasaki K., Hashimoto H., Makino S.;
RT "Molecular cloning and complete nucleotide sequence of genomic RNA of
RT the AIK-C strain of attenuated measles virus.";
RL Virus Genes 7:67-81(1993).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S58435; AAB26145.1; -.
CC DR PIR; E48556; E48556.
CC DR HSP; P04849; 1SVF.
CC DR InterPro; IPR000776; Fusion_gly.
CC Pfam; PF00523; fusion_gly; 1.
CC KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
CC SIGNAL 1 23
CC FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
CC FT CHAIN 24 112 PROTEIN F2.
CC FT CHAIN 113 550 PROTEIN F1.
CC FT TRANSMEM 137 494 POTENTIAL.
CC FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 495 515 POTENTIAL.
CC FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
CC FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 550 AA; 59540 MW; AAC4DAB92DE0D938 CRC64;

Query Match 41.4%; Score 72; DB 1; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.0082;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 4
VGLF_MEASE STANDARD; PRT; 550 AA.
AC P08300;
```

DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
 DE Fusion glycoprotein F1].
 GN F.
 OS Measles virus (strain Edmonston) (Subacute sclerosing panencephalitis
 OS virus).
 OS Measles virus (strain Halle) (Subacute sclerosing panencephalitis
 OS virus).
 OS Measles virus (strain Leningrad-16) (Subacute sclerosing panencephalitis
 OS virus).
 OS Measles virus (strain Edmonston-Zagreb) (Subacute sclerosing
 OS panencephalitis virus).
 OS Measles virus (strain Philadelphia-26) (Subacute sclerosing
 OS panencephalitis virus), and
 OS Measles virus (strain Edmonston B) (Subacute sclerosing panencephalitis
 OS virus).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
 OC NCBI_TaxID=11235, 11236, 70147, 70149, 70148, 70146;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EDMONSTON;
 RX MEDLINE=87071668; PubMed=3788062;
 RA Richardson C.D., Hull D., Greer P., Hasel K., Berkovich A.,
 RA Englund G., Bellini W.J., Rima B., Lazzarini R.A.;
 RA "The nucleotide sequence of the mRNA encoding the fusion protein of
 RT measles virus (Edmonston strain): a comparison of fusion proteins
 RT from several different paramyxoviruses.";
 RL Virology 155:508-523(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HALLE;
 RX MEDLINE=87224816; PubMed=3585281;
 RA Buckland R., Gerald C., Barker R., Wild T.F.;
 RA "Fusion glycoprotein of measles virus: nucleotide sequence of the
 RT gene and comparison with other paramyxoviruses.";
 RL J. Gen. Virol. 68:1695-1703(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EDMONSTON;
 RX MEDLINE=90085790; PubMed=2596022;
 RA Cattaneo R., Schmid A., Spielhofer P., Kaelin K., Bacsko K.,
 RA Meulen V., Pardowitz J., Flanagan S., Rima B.K., Udem S.A.;
 RA "Mutated and hypermutated genes of persistent measles viruses which
 RT caused lethal human brain diseases.";
 RL Virology 173:415-425(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EDMONSTON;
 RX MEDLINE=92263801; PubMed=1585658;
 RA Schmid A., Spielhofer P., Cattaneo R., Bacsko K., Ter Meulen V.,
 RA Billeter M.A.;
 RA "Subacute sclerosing panencephalitis is typically characterized by
 RT alterations in the fusion protein cytoplasmic domain of the
 RT persisting measles virus.";
 RL Virology 188:910-915(1992).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EDMONSTON, LENINGRAD-16, AND EDMONSTON-ZAGREB;
 RX MEDLINE=94249283; PubMed=8191786;
 RA Rota J.S., Wang Z.D., Rota P.A., Bellini W.J.;
 RA "Comparison of sequences of the H, F, and N coding genes of measles
 RT virus vaccine strains.";
 RL Virus Res. 31:317-330(1994).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PHILADELPHIA-26;
 RX MEDLINE=94303181; PubMed=8030232;
 RA Hummel K.B., Vanchiere J.A., Bellini W.J.;
 RA "Restriction of fusion protein mRNA as a mechanism of measles virus
 RT persistence.";

RL Virology 202:665-672(1994).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EDMONSTON B;
 RA Billeter M.A.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
 CC MEMBRANES.
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
 CC LINKED BY A DISULFIDE BOND.
 CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M14915; AAA46423.1; -;
 DR EMBL; X05597; CAA29090.1; ALT_INIT.
 DR EMBL; K01711; AAA75498.1; ALT_INIT.
 DR EMBL; K01711; AAA75499.1; -;
 DR EMBL; U03657; AAA56647.1; ALT_INIT.
 DR EMBL; U03659; AAA56649.1; ALT_INIT.
 DR EMBL; U03670; AAA56660.1; ALT_INIT.
 DR EMBL; U08416; AAA50550.1; ALT_INIT.
 DR EMBL; Z66517; CAA91367.1; ALT_INIT.
 DR EMBL; Z66517; CAA91368.1; -;
 DR PIR; A26962; VGN2MV.
 DR HSP; P04849; ISVF.
 DR InterPro; IPR000776; Fusion_gly.
 DR Pfam; PF00523; fusion_gly; 1.
 KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 550 FUSION GLYCOPROTEIN F0.
 FT CHAIN 24 112 PROTEIN F2.
 FT CHAIN 113 550 POTENTIAL.
 FT TRANSMEM 113 136 POTENTIAL.
 FT DOMAIN 137 494 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 495 515 POTENTIAL.
 FT DOMAIN 516 550 CYTOPLASMIC (POTENTIAL).
 FT DISULFD 68 195 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
 FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 550 AA; 59532 MW; 7AA4F1CA82169093 CRC64;
 Query Match 41.4%; Score 72; DB 1; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.0082;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 LSEIKGVIVHRLEGV 33
 Db 288 LSEIKGVIVHRLEGV 302
 RESULT 5
 VGLF_RINDR
 ID VGLF_RINDR STANDARD; PRT; 546 AA.
 AC P41356;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
 DE Fusion glycoprotein F1].
 GN F.
 OS Rinderpest virus (strain RBOK) (RDV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
 OC NCBI_TaxID=36409;

```
RN SEQUENCE FROM N.A.
RX MEDLINE=95088609; PubMed=7996154;
RA Evans S.A., Baron M.D., Chamberlain R.W., Goatley L., Barrett T.;
RT "Nucleotide sequence comparisons of the fusion protein gene from
RT virulent and attenuated strains of rinderpest virus.";
RL J. Gen. Virol. 75:3611-3617(1994).
CC
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z30700; CAA83186.1; -
DR EMBL; Z30697; CAA83181.1; -
DR PIR; S47305; S47305.
DR HSP; P04849; LSVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58705 MW; ED3DFAFFDBECB95 CRC64;

Query Match 40.8%; Score 71; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.011;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
| | | | | | | | | | | | | | |
Db 284 LSEIKGVIVHRLEGV 298

RESULT 6
VGLF_RINDL STANDARD; PRT; 546 AA.
AC P10864;
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain L) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11243;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=88219541; PubMed=3285575;
RA Tsukiyama K., Yoshikawa Y., Yamanouchi K.;
RT "Fusion glycoprotein (F) of rinderpest virus: entire nucleotide
RT sequence of the F mRNA, and several features of the F protein.";
```

```
RL Virology 164:523-530(1988).
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20870; AAA47399.1; -
DR PIR; A28921; VGNZRL.
DR HSP; P04849; LSVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
FT CHAIN 20 108 F2 PROTEIN.
FT CHAIN 109 546 F1 PROTEIN.
FT DOMAIN 104 108 ARG/LYS-RICH (BASIC).
FT TRANSMEM 109 133 POTENTIAL.
FT TRANSMEM 484 513 POTENTIAL.
FT DOMAIN 514 517 ARG/LYS-RICH (BASIC).
FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 546 AA; 58911 MW; 985029418F28FFB5 CRC64;

Query Match 37.9%; Score 66; DB 1; Length 546;
Best Local Similarity 93.3%; Pred. No. 0.063;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 19 LSEIKGVIVHRLEGV 33
| | | | | | | | | | | | | | |
Db 284 LSEIKGVIVHRLEGV 298

RESULT 7
VGLF_CDVO STANDARD; PRT; 662 AA.
AC P12569; Q65991;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11233;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=86129050; PubMed=3433924;
RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;
RT "The nucleotide sequence of the gene encoding the F protein of canine
RT distemper virus: a comparison of the deduced amino acid sequence with
RT other paramyxoviruses.";
RL Virus Res. 8:373-386(1987).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=93227696; PubMed=8470428;
RA Wild T.F., Bernard A., Spehner D., Villevall D., Drillien R.;
RT "Vaccination of mice against canine distemper virus-induced
RT encephalitis with vaccinia virus recombinants encoding measles or
```

```
RT canine distemper virus antigens." ;
RL Vaccine 1:438-444(1993).
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

EMBL; M21849; AAA42878.1; -;
DR EMBL; M65509; CAA46481.1; -;
DR PIR; J50321; VGNZCD.
DR HSP; P04849; LSVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
FT SIGNAL 1 662
FT CHAIN ? 662 FUSION GLYCOPROTEIN F0.
FT CHAIN ? 224 PROTEIN F2.
FT CHAIN 225 662 PROTEIN F1.
FT TRANSMEM 606 629 POTENTIAL.
FT DISULFID 180 307 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 62 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 3 3 R -> K (IN REF. 2).
FT CONFLICT 140 140 D -> N (IN REF. 2).
FT CONFLICT 152 152 N -> S (IN REF. 2).
FT CONFLICT 171 171 I -> M (IN REF. 2).
FT CONFLICT 174 174 A -> V (IN REF. 2).
FT CONFLICT 562 562 L -> H (IN REF. 2).
SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;

Query Match 37.9%; Score 66; DB 1; Length 662;
Best Local Similarity 73.7%; Pred. No. 0.078;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 15 SKLLSEIKGVIVHRLEGV 33
I III:IIIIIIII I
Db 396 SYPTLSEVKGVIVHRLEAV 414

RESULT 8
GONI_SPAU
ID GONI_SPAU STANDARD; PRT; 95 AA.
AC P51919;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadotropin-releasing hormone I (GnRH-I)
DE (LH-RH I) (Lubiberin I) (SEGNRH).
GN GNRH1.
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoldei;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=952684499; PubMed=7749463;
RA Gotthilf Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-

```
RT releasing hormone from the gilthead seabream (Sparus aurata).";  
RL Mol. Mar. Biol. Biotechnol. 4:27-35(1995).  
RN [2]  
RP SEQUENCE OF 26-35.  
RC TISSUE=Brain;  
RX MEDLINE=95083645; PubMed=7991588;  
RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,  
RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;  
RT "Three forms of gonadotropin-releasing hormone characterized from  
RT brains of one species.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).  
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.  
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

EMBL; U30320; AAA75469.1; -;
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
FT SIGNAL; Multigene family.
FT SIGNAL 1 25
FT CHAIN 26 95 PRONADOLIBERIN I.
FT PEPTIDE 26 35 GONADOLIBERIN I.
FT PEPTIDE 39 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 35 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
SQ SEQUENCE 95 AA; 10753 MW; 49313FD6FD6B87DA CRC64;

Query Match 37.6%; Score 65.5; DB 1; Length 95;
Best Local Similarity 45.5%; Pred. No. 0.01;
Matches 15; Conservative 3; Mismatches 14; Indels 1; Gaps 1;

Qy 2 HWSYGLRPGSGPSLKLSEIKGVIVHRLEGE 34
IIIIIIII II II: I I I I I I
Db 27 HWSYGLSPGKG-RDLDSLSLTGLNIIRPPHYD 58

RESULT 9
VGLF_RINDK
ID VGLF_RINDK STANDARD; PRT; 546 AA.
AC P12574;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Rinderpest virus (strain Kabete O) (RDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses;
OX NCBI_TaxID=11242;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88322864; PubMed=3413983;
RA Hsu D., Yamanaka M., Miller J., Dale B., Grubman M., Yilma T.;
RT "Cloning of the fusion gene of rinderpest virus: comparative sequence
RT analysis with other morbilliviruses.";
RL Virology 166:149-153(1988).
CC -!- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN


```

RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RL Neuroendocrinology 60:346-359(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S75918; AAB33096.1; -.
DR InterPro; IPR002012; GNRH.
DR PIR; A01410; RHUG.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 5
FT CHAIN 6 >67 PROGNADOLIBERIN I.
FT PEPTIDE 6 15 GONADOLIBERIN I.
FT ACT_SITE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 6 6 ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 15 15 SIMILARITY).
FT NON_TER 67 67
FT SEQUENCE 67 AA; 7573 MW; 505394DAA361A3F2 CRC64;
Query Match 36.8%; Score 64; DB 1; Length 67;
Best Local Similarity 40.0%; Pred. No. 0.012;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSGPSLKLSEIKGVI 26
Db 7 HWSYGLRPGGKRDAENLMDSPQEIV 31

RESULT 12
GONI_HUMAN
ID GONI_HUMAN STANDARD; PRT; 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated
DE peptide I].
GN GNRH1 OR GNRH OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89366682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RT hormone gene.";
RL Nucleic Acids Res. 17:6403-6403(1989).

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone.";
RL Nature 311:666-668(1984).
RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE=83126573; PubMed=6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
RT biosynthesized in the human placenta.";
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC Lutrepulse or Lutrelaf (Ferring Pharmaceuticals) and Relisorm
CC (Serono).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01059; CAA25526.1; -.
DR EMBL; M12578; AAA35916.1; -.
DR EMBL; X15215; CAA33285.1; -.
DR PIR; A01410; RHUG.
DR PIR; A26173; A26173.
DR PIR; S05308; S05308.
DR MIN; 152760; -.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBRN1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 24 24 ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT CONFLICT 16 16 W -> S (IN REF. 3).
FT SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;
Query Match 36.8%; Score 64; DB 1; Length 92;
Best Local Similarity 40.0%; Pred. No. 0.017;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSGPSLKLSEIKGVI 26
Db 25 HWSYGLRPGGKRDAENLMDSPQEIV 49

RESULT 13

```

VGLEF_MEASI STANDARD; PRT; 529 AA.
AC P26031; O83298;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F
OS Measles virus (strain IP-3-Ca) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
ON NCBI_TaxID=11237;
RX MEDLINE=92263801; PubMed=1585658;
RP Billeter M.A.;
RA Schmid A., Spielhofer P., Cattaneo R., Bacsko K., Ter Meulen V.,
RA "Subacute sclerosing panencephalitis is typically characterized by
alterations in the fusion protein cytoplasmic domain of the
FT persisting measles virus";
FT Virology 188:910-915(1992).
RL
CC -1- FUNCTION: THIS PROTEIN DIRECTS FUSION OF VIRAL AND CELLULAR
CC MEMBRANES.
CC -1- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES FUSION GLYCOPROTEIN
CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X16566; CAA34567.1; -;
DR EMBL; X16566; CAA34568.1; ALT_INIT.
DR HSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_protein; Transmembrane; Envelope protein; Signal.
KW Glycoprotein; Fusion protein; Fusion_gly; 1.
FT SIGNAL 1 26
FT CHAIN 27 529 BY SIMILARITY.
FT CHAIN 27 115 FUSION GLYCOPROTEIN F0.
FT CHAIN 116 529 PROTEIN F2.
FT CHAIN 116 529 PROTEIN F1.
FT TRANSMEM 116 139 POTENTIAL.
FT DOMAIN 140 497 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 498 518 POTENTIAL.
FT DOMAIN 519 529 CYTOPLASMIC (POTENTIAL).
FT DISULFD 71 198 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 529 AA; 57331 MW; A5987BC9F07E9AA9 CRC64;

Query Match 36.8%; Score 64; DB 1; Length 529;
Best Local Similarity 93.3%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 291 LSEIKGVIVHRLGV 305

RESULT 14
GON1_PAGMA
ID GON1_PAGMA STANDARD; PRT; 95 AA.
AC P70074;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
DE (LH-RH I) (Luliberin I).
GN GNRH1.
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Pagrus.
ON NCBI_TaxID=143350;
RX MEDLINE=143350;
RP SEQUENCE FROM N.A.
RA TISSUE=Brain;
RA Okuzawa K., Granneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D86582; BAA13129.1; -;
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 24 95 POTENTIAL.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 33 33 SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP)
FT SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;

Query Match 35.9%; Score 62.5; DB 1; Length 95;
Best Local Similarity 50.0%; Pred. No. 0.029;
Matches 14; Conservative 2; Mismatches 11; Indels 1; Gaps 1;

QY 2 HWSYGLRGSSGSLKLLSEIKGVIVHR 29
DB 25 HWSYGLSPGGK-RDLDSLTGLGDIIR 51

RESULT 15
GON1_XENLA
ID GON1_XENLA STANDARD; PRT; 89 AA.
AC P45656;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-NOV-2000 (Rel. 39, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
DE (LH-RH) (Luliberin I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
ON NCBI_TaxID=8355;
RX MEDLINE=8355;
RP SEQUENCE FROM N.A.
RA TISSUE=Forebrain;
RA MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Wray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a

```

RT mammalian-like expression pattern and conserved domains in
RT GnRH-associated peptide, but brain onset is delayed until
RT metamorphosis.;
RL Endocrinology 134:1835-1844(1994).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28040; AAA49728.1; -.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 89
FT PEPTIDE 24 33
FT PEPTIDE 37 89
FT PEPTIDE 37 85
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 33 33
SQ SEQUENCE 89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;
Query Match 35.6%; Score 62; DB 1; Length 89;
Best Local Similarity 25.4%; Pred. No. 0.031;
Matches 15; Conservative 7; Mismatches 9; Indels 28; Gaps 1;
QY 2 HWSYGLRPGSSGPSLKL-----LSEIKGVIVHRLEG 32
Db 25 HWSYGLRPGGRDTESLQDMYHETPNEVALFPELEPCSVPSRLNVLRLGALMNWLEG 83

```

Search completed: October 10, 2002, 16:06:32
Job time : 5.43162 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 13.2222 Seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-13
Perfect score: 174
Sequence: 1 XHNSYGLRPGSSGPSLKLSEIKGIVVHRLEGV 34

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	41.4	534	12	Q04243 measles vir
2	72	41.4	537	12	Q04242 measles vir
3	72	41.4	545	12	Q9PXA4 measles vir
4	72	41.4	546	12	Q91HA5 rinderpest
5	72	41.4	550	12	Q89495 measles vir
6	72	41.4	550	12	P90331 measles vir
7	72	41.4	550	12	Q9YJ94 measles vir
8	72	41.4	550	12	Q9QEX1 measles vir
9	72	41.4	550	12	Q9QEX0 measles vir
10	72	41.4	550	12	Q9QEW9 measles vir
11	72	41.4	550	12	Q9QEW8 measles vir
12	72	41.4	550	12	Q9QEW7 measles vir
13	72	41.4	550	12	Q9WMK4 measles vir
14	72	41.4	550	12	P90330 measles vir
15	72	41.4	553	12	Q93055 measles vir
16	72	41.4	553	12	Q83530 measles vir

17	72	41.4	553	12	O11383	O11383 measles vir
18	72	41.4	553	12	Q83518	Q83518 measles vir
19	72	41.4	553	12	Q83521	Q83521 measles vir
20	72	41.4	553	12	Q83525	Q83525 measles vir
21	72	41.4	553	12	Q83527	Q83527 measles vir
22	72	41.4	553	12	Q83533	Q83533 measles vir
23	72	41.4	553	12	Q83536	Q83536 measles vir
24	72	41.4	553	12	Q9IC36	Q9IC36 measles vir
25	72	41.4	553	12	Q9IFK2	Q9IFK2 measles vir
26	72	41.4	553	12	P88973	P88973 measles vir
27	72	41.4	553	12	P88974	P88974 measles vir
28	72	41.4	553	12	O91248	O91248 measles vir
29	72	41.4	553	12	O04244	O04244 measles vir
30	72	41.4	553	12	Q91QP2	Q91QP2 measles vir
31	72	41.4	579	12	Q9PW04	Q9PW04 measles vir
32	68	39.1	545	12	Q9QEW6	Q9QEW6 measles vir
33	68	39.1	553	12	O11380	O11380 measles vir
34	66	37.9	528	12	O9YJW9	O9YJW9 canine dist
35	66	37.9	662	12	O9YK17	O9YK17 canine dist
36	66	37.9	662	12	O89327	O89327 canine dist
37	66	37.9	662	12	Q9DX22	Q9DX22 canine dist
38	66	37.9	662	12	Q91KN3	Q91KN3 canine dist
39	65.5	37.6	87	13	O9YI26	O9YI26 sparus aura
40	65	37.4	552	12	O66147	O66147 cetacean mo
41	62	35.6	552	12	O56852	O56852 dolphin mor
42	62	35.6	552	12	O66409	O66409 dolphin mor
43	61	35.1	91	13	Q9PRH0	Q9PRH0 anguilla ja
44	61	35.1	553	12	Q83629	Q83629 measles vir
45	60	34.5	546	12	Q84926	Q84926 peste-des-p

ALIGNMENTS

RESULT 1

Q04243
ID Q04243 PRELIMINARY; PRT; 534 AA.
AC Q04243;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Baczko K., ter Meulen V.,
RA Billetter M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RL Cell 55:255-265(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billetter M.A.;
RL Virology 0:0-0(0).
DR EMBL; X16568; CAA34581.1; -.
DR EMBL; X16568; CAA34582.1; -.
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 534 AA: 57899 MW; 637245E23B5BE044 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 534;

Best Local Similarity 100.0%; Pred. No. 0.062;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33

Db 291 LSEIKGVIVHRLEGV 305

```

RESULT 2
Q04242 Q04242 PRELIMINARY; PRT; 537 AA.
ID AC Q04242;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003063; PubMed=3167982;
RA Cattaneo R., Schmid A., Eschle D., Bacsko K., ter Meulen V.,
RA Billette M.A.;
RT "Biased hypermutation and other genetic changes in defective measles
RT viruses in human brain infections.";
RL Cell 55:255-265(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Cattaneo R., Billette M.A.;
RL Virology 0:0-0(0).
DR EMBL; X16567; CRA34574.1; -.
DR EMBL; X16567; CRA34575.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 537 AA; 58275 MW; DOA60AC66D979E06 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 537;
Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 291 LSEIKGVIVHRLGV 305

RESULT 3
Q9PXA4 Q9PXA4 PRELIMINARY; PRT; 545 AA.
ID AC Q9PXA4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=OSA-3;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179440; AAF02705.1; -.
DR EMBL; AF179439; AAF02704.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 545 AA; 58907 MW; 0234C28AE193E77D CRC64;

Query Match 41.4%; Score 72; DB 12; Length 545;

```

```

Best Local Similarity 100.0%; Pred. No. 0.063;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 288 LSEIKGVIVHRLGV 302

RESULT 4
Q91HA5 Q91HA5 PRELIMINARY; PRT; 546 AA.
ID AC Q91HA5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Rinderpest virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11241;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K; PubMed=11186456;
RX MEDLINE=21014265;
RA Aianot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA Gusev A.A.;
RT "Primary structure of the F-gene from Rinderpest virus strain K.";
RL Mol. Gen. Microbiol. Virusol. 4:29-33(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K;
RA Ayanot P.K., Sminev A.G., Bezborodova S.V., Starov S.K., Drygin V.V.,
RA Gusev A.A.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035887; AAK63190.1; -.
SQ SEQUENCE 546 AA; 58572 MW; 449B2BDD7405F0B CRC64;

Query Match 41.4%; Score 72; DB 12; Length 546;
Best Local Similarity 100.0%; Pred. No. 0.063;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGV 33
DB 284 LSEIKGVIVHRLGV 298

RESULT 5
Q89495 Q89495 PRELIMINARY; PRT; 550 AA.
ID AC Q89495;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92230209; PubMed=1566568;
RA Rota J.S., Hummel K.B., Rota P.A., Bellini W.J.;
RT "Genetic variability of the glycoprotein genes of current wild-type
RT measles isolates.";
RL Virology 188:135-142(1992).
DR EMBL; M81903; AAA46422.1; -.
DR EMBL; M81901; AAA46421.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59564 MW; A78EEC9CD6268E58 CRC64;

```

Query Match 41.4%; Score 72; DB 12; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.064;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 DB 288 LSEIKGVIVHRLEGV 302

RESULT 6

P90331 ID P90331 PRELIMINARY; PRT; 550 AA.
 AC P90331.
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE FUSION PROTEIN.
 GN F.
 OS Measles virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
 OX NCBI_TaxID=11234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NAGAHATA;
 RA Sheng J., Watanabe M., Ueda S.;
 RA "Selection of a neurotropic variant of measles virus.";
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NAGAHATA;
 RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;
 RA "An amino acid alteration of F protein responsible for the enhanced
 fusogenicity of measles virus.";
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NAGAHATA;
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
 RA Ogura H.;
 RA "Nucleotide sequences of the fusion protein gene of subacute
 sclerosing panencephalitis viruses: deduced amino acid sequences
 showed the cytoplasmic domain highly mutated --truncated, elongated or
 RT predicted secondary structure changed.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D63926; BAA0958.1; -;
 DR EMBL; AF179431; AAF02696.1; -;
 DR HSSP; P04849; ISVF.
 DR InterPro; IPR000776; Fusion_gly.
 DR Pfam; PF00523; fusion_gly; 1.
 SQ SEQUENCE 550 AA; 59530 MW; 97C991C7E2169839 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.064;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 DB 288 LSEIKGVIVHRLEGV 302

RESULT 7

Q9YJ94 ID Q9YJ94 PRELIMINARY; PRT; 550 AA.
 AC Q9YJ94;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE FUSION PROTEIN.
 OS Measles virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.

OX NCBI_TaxID=11234;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-9301V;
 RX MEDLINE=98440529; PubMed=9765410;
 RA Takeda M., Kato A., Kobune F., Sakata H., Li Y., Shioda T., Sakai Y.,
 RA Asakawa M., Nagai Y.;
 RA "Measles virus attenuation associated with transcriptional impediment
 RT and a few amino acid changes in the polymerase and accessory
 RT proteins.";
 RL J. Virol. 72:8690-8696(1998).
 DR EMBL; AB012949; BAA33877.1; -;
 DR EMBL; AB012948; BAA33871.1; -;
 DR HSSP; P04849; ISVF.
 DR InterPro; IPR000776; Fusion_gly.
 DR Pfam; PF00523; fusion_gly; 1.
 SQ SEQUENCE 550 AA; 59512 MW; 7AA4F1D117197BF9 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.064;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 DB 288 LSEIKGVIVHRLEGV 302

RESULT 8

Q9QEX1 ID Q9QEX1 PRELIMINARY; PRT; 550 AA.
 AC Q9QEX1.
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE FUSION PROTEIN.
 OS Measles virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
 OX NCBI_TaxID=11234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MASUSAKO;
 RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
 RA Ogura H.;
 RA "Nucleotide sequences of the fusion protein gene of subacute
 sclerosing panencephalitis viruses: deduced amino acid sequences
 RT showed the cytoplasmic domain highly mutated --truncated, elongated or
 RT predicted secondary structure changed.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF179430; AAF02695.1; -;
 DR HSSP; P04849; ISVF.
 DR InterPro; IPR000776; Fusion_gly.
 DR Pfam; PF00523; fusion_gly; 1.
 SQ SEQUENCE 550 AA; 59559 MW; 609EE024A7E59C54 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.064;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
 DB 288 LSEIKGVIVHRLEGV 302

RESULT 9

Q9QEX0 ID Q9QEX0 PRELIMINARY; PRT; 550 AA.
 AC Q9QEX0.
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE FUSION PROTEIN.
 OS Measles virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=TOYOSHIMA;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179432; AAF02697.1; -
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59504 MW; 2AA969D37FA5CA17 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 10

Q9QEW9 PRELIMINARY; PRT; 550 AA.
ID Q9QEW9;
AC Q9QEW9;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179436; AAF02701.1; -
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59405 MW; 0AF6DBFC5DD22BBA CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 11

Q9QEW8 PRELIMINARY; PRT; 550 AA.
ID Q9QEW8;
AC Q9QEW8;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE FUSION PROTEIN.

OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179437; AAF02702.1; -
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59315 MW; 086E51FED235EBBA CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 12

Q9QEW7 PRELIMINARY; PRT; 550 AA.
ID Q9QEW7;
AC Q9QEW7;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=OSA-2;
RA Ning X., Ayata M., Morimoto K., Ito N., Shingai M., Kimura M.,
RA Ogura H.;
RT "Nucleotide sequences of the fusion protein gene of subacute
RT sclerosing panencephalitis viruses: deduced amino acid sequences
RT showed the cytoplasmic domain highly mutated --truncated, elongated or
RT predicted secondary structure changed.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179438; AAF02703.1; -
DR HSSP; P04849; 1SVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59333 MW; 086E51FED5582BBA CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
|||||
Db 288 LSEIKGVIVHRLEGV 302

RESULT 13

Q9WMK4 PRELIMINARY; PRT; 550 AA.
ID Q9WMK4;
AC Q9WMK4;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)


```
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WTF;
RX MEDLINE=99329215; PubMed=10400788;
RA Johnston I.C., Ter Meulen V., Schneider-Schaulies J.,
RA Schneider-Schaulies S.;
RT "A recombinant measles vaccine virus expressing wild-type
RT glycoproteins : consequences for viral spread and cell tropism.";
RL J. Virol. 73:6903-6915(1999).
DR EMBL; AJ133108; CAB38075.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59580 MW; 8255499968B5D862 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302
|||||

RESULT 14
P90330
ID P90330 PRELIMINARY; PRT; 550 AA.
AC P90330;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FUSION PROTEIN.
GN F.
OS Measles virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA(HB);
RA Sheng J., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced
RT fusogenicity of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGAHATA(HB);
RA Sheng J., Nakanishi M., Watanabe M., Ueda S.;
RT "An amino acid alteration of F protein responsible for the enhanced
RT fusogenicity of measles virus.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; D63924; BAA09951.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
SQ SEQUENCE 550 AA; 59589 MW; 73E7BD457ABA39B7 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

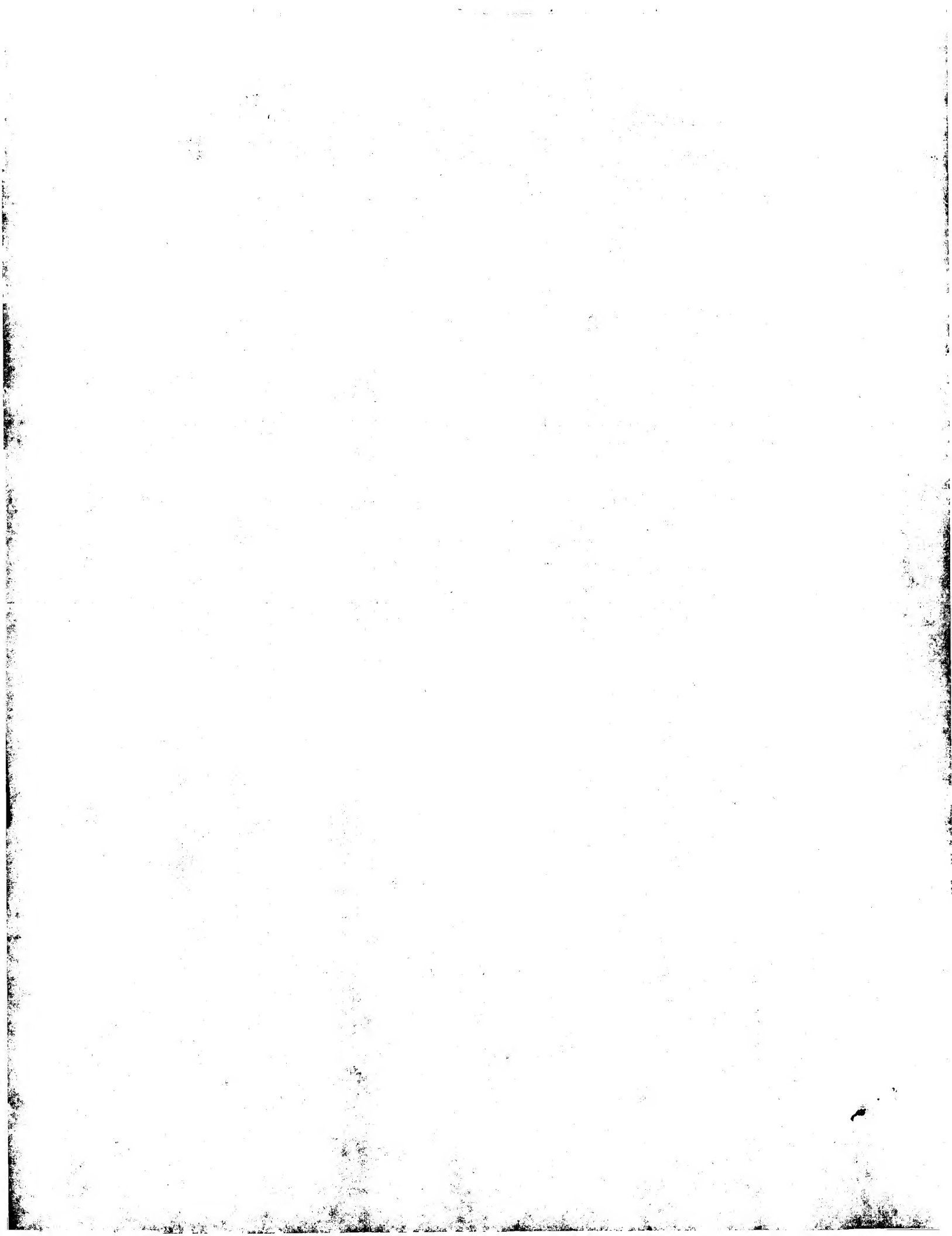
QY 19 LSEIKGVIVHRLEGV 33
Db 288 LSEIKGVIVHRLEGV 302
|||||

RESULT 15
O93055
ID O93055 PRELIMINARY; PRT; 553 AA.
AC O93055;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FUSION GLYCOPROTEIN PRECURSOR (FUSION PROTEIN).
GN F.
OS Measles virus, and
OS Measles virus (strain Edmonston-Zagreb) (Subacute sclerose
OS panencephalitis virus).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11234, 70149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HALLE;
RX MEDLINE=87224816; PubMed=3585281;
RA Buckland R., Gerald C., Barker R., Wild T.F.;
RT "Fusion glycoprotein of measles virus: nucleotide sequence of the gene
RT and comparison with other paramyxoviruses.";
RL J. Gen. Virol. 68:1695-1703(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Parks C.L., Lerch R.A., Walpita P., Wang H.-P., Sidhu M.S., Udem S.A.;
RT "Comparison of predicted amino acid sequences from measles virus
RT strains in the Edmonston vaccine lineage.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; D00090; BAA00056.1; -.
DR EMBL; AF266290; AAF85696.1; -.
DR EMBL; AF266288; AAF85680.1; -.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; fusion_gly; 1.
KW Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 115 FUSION GLYCOPROTEIN F2 SUBUNIT.
FT CHAIN 116 553 FUSION GLYCOPROTEIN F1 SUBUNIT.
SQ SEQUENCE 553 AA; 59863 MW; 94C616494DDF8023 CRC64;

Query Match 41.4%; Score 72; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLEGV 33
Db 291 LSEIKGVIVHRLEGV 305
|||||

Search completed: October 10, 2002, 16:10:00
Job time : 14.2222 secs
```



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 6.98077 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKIAKMEKASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	134.5	79.6	33	1	US-08-446-692-27
2	134.5	79.6	33	2	US-08-488-351A-27
3	89	52.7	21	1	US-08-186-266-6
4	89	52.7	21	1	US-08-446-692-48
5	89	52.7	21	2	US-08-488-351A-48
6	89	52.7	21	3	US-09-100-409A-54
7	89	52.7	21	5	PCT-US95-02121-97
8	89	52.7	21	5	PCT-US95-13841-20
9	86	50.9	423	2	US-08-760-797A-1
10	86	50.9	424	2	US-08-760-797A-3
11	86	50.9	424	4	US-08-932-929B-1
12	86	50.9	424	4	US-08-932-929B-3
13	82	48.5	412	1	US-08-313-288B-18
14	79	46.7	17	5	PCT-US95-02121-96
15	75	44.4	16	2	US-08-817-933A-7
16	73.5	43.5	20	1	US-08-465-167A-20
17	73.5	43.5	20	5	PCT-US92-07218-17
18	70	41.4	15	6	5169933-30
19	66	39.1	30	3	US-09-100-414B-71
20	66	39.1	30	4	US-09-303-323-71
21	65	38.5	30	3	US-09-100-414B-73
22	65	38.5	30	4	US-09-303-323-73
23	63.5	37.6	16	1	US-08-485-167A-19
24	63.5	37.6	16	5	PCT-US92-07218-16
25	62	36.7	21	1	US-08-305-871A-12
26	61	36.1	30	3	US-09-100-414B-68
27	61	36.1	30	4	US-09-303-323-68

```

28 59 34.9 49 1 US-08-387-156-4 Sequence 4, Appli
29 59 34.9 49 2 US-08-694-863-4 Sequence 4, Appli
30 59 34.9 49 2 US-08-878-748-4 Sequence 4, Appli
31 59 34.9 49 3 US-09-124-491-4 Sequence 4, Appli
32 59 34.9 544 1 US-08-387-156-10 Sequence 10, Appl
33 59 34.9 544 2 US-08-694-863-10 Sequence 10, Appl
34 59 34.9 544 2 US-08-878-748-10 Sequence 10, Appl
35 59 34.9 544 3 US-09-124-491-10 Sequence 10, Appl
36 59 34.9 699 2 US-08-694-863-16 Sequence 16, Appl
37 59 34.9 699 3 US-09-124-491-16 Sequence 16, Appl
38 59 34.9 977 1 US-08-387-156-8 Sequence 8, Appli
39 59 34.9 977 2 US-08-694-863-8 Sequence 8, Appli
40 59 34.9 977 2 US-08-878-748-8 Sequence 8, Appli
41 59 34.9 977 3 US-09-124-491-8 Sequence 8, Appli
42 58 34.3 31 1 US-08-446-692-30 Sequence 30, Appl
43 58 34.3 31 2 US-08-488-351A-30 Sequence 30, Appl
44 58 34.3 47 1 US-08-446-692-35 Sequence 35, Appl
45 58 34.3 47 2 US-08-488-351A-35 Sequence 35, Appl

```

ALIGNMENTS

```

RESULT 1
US-08-446-692-27
; Sequence 27, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Y1
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-27

```

```

Query Match 79.6%; Score 134.5; DB 1; Length 33;
Best Local Similarity 90.3%; Pred. No. 1.8e-13;
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 2 EKIAKMEKASSVFNVNSGSLHWSYGLRP 32
Db 3 EKIAKMEKASSVFNVNSGGE-HWSYGLRP 32

```

```
RESULT 2
US-08-488-351A-27
; Sequence 27, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-27

Query Match 79.6%; Score 134.5; DB 2; Length 33;
Best Local Similarity 90.3%; Pred. No. 1.8e-13;
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 EKIKAKMEKASSVFNVNSGSLHWSYGLRP 32
Db 3 EKIKAKMEKASSVFNVNSGGE-HWSYGLRP 32

RESULT 3
US-08-186-266-6
; Sequence 6, Application US/08186266
; Patent No. 5662907
; GENERAL INFORMATION:
; APPLICANT: KUBO, Ralph T.
; APPLICANT: GREY, Howard M.
; APPLICANT: SETTE, Alessandro
; APPLICANT: CELIS, Esteban
```

```
; TITLE OF INVENTION: INDUCTION OF ANTI-TUMOR CYTOTOXIC
; TITLE OF INVENTION: T LYMPHOCYTES IN HUMANS USING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,266
; FILING DATE: 25-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/159,339
; FILING DATE: 29-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-50-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Plasmodium falciparum CS
; OTHER INFORMATION: protein at positions 378-398."
US-08-186-266-6

Query Match 52.7%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKIKAKMEKASSVFNVNS 20
Db 3 EKIKAKMEKASSVFNVNS 21

RESULT 4
US-08-446-692-48
; Sequence 48, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Maria C.H. Lin
 STREET: 345 Park Avenue
 CITY: New York
 STATE: NY
 COUNTRY: US
 ZIP: 10154-0053
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,692
 FILING DATE: 7-JUN-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Maria C.H. Lin
 REGISTRATION NUMBER: 29,323
 REFERENCE/DOCKET NUMBER: 1151-4146 US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)415-8745
 TELEFAX: (516)751-6849
 INFORMATION FOR SEQ ID NO: 48:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 21 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-446-692-48

Query Match 52.7%; Score 89; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 6.5e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKTIKMEKASSVFNVNS 20
 Db 3 EKTIKMEKASSVFNVNS 21

RESULT 5

US-08-488-351A-48
 Sequence 48, Application US/08488351A
 Patent No. 5843446
 GENERAL INFORMATION:
 APPLICANT: Ladd, Anna
 APPLICANT: Wang, Chang Yi
 APPLICANT: Zamb, Timothy
 TITLE OF INVENTION: Immunogenic LHRH peptide constructs
 TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
 NUMBER OF SEQUENCES: 114
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Maria C.H. Lin
 STREET: 345 Park Avenue
 CITY: New York
 STATE: NY
 COUNTRY: US
 ZIP: 10154-0053
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,692
 FILING DATE: 7-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/446,692
 FILING DATE: 7-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/229,275
 FILING DATE: 14-APR-1994

CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/057,166
 FILING DATE: 27-APR-1992
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Maria C.H. Lin
 REGISTRATION NUMBER: 29,323
 REFERENCE/DOCKET NUMBER: 1151-4146 US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)415-8745
 TELEFAX: (516)751-6849
 INFORMATION FOR SEQ ID NO: 48:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 21 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-488-351A-48

Query Match 52.7%; Score 89; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 6.5e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKTIKMEKASSVFNVNS 20
 Db 3 EKTIKMEKASSVFNVNS 21

RESULT 6

US-09-100-409A-54
 Sequence 54, Application US/09100409A
 Patent No. 6090388
 GENERAL INFORMATION:
 APPLICANT: Wang, Chang Yi
 TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
 TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
 TITLE OF INVENTION: IMMUNE DISORDERS
 NUMBER OF SEQUENCES: 64
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 Park Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10154-0054
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version
 SOFTWARE: #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/100,409A
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME:
 REGISTRATION NUMBER:
 REFERENCE/DOCKET NUMBER: 1151-4154
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-758-4800
 TELEFAX: 212-751-6849
 INFORMATION FOR SEQ ID NO: 54:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 21 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-100-409A-54

Query Match 52.7%; Score 89; DB 3; Length 21;
 Best Local Similarity 100.0%; Pred. No. 6.5e-07;

```
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIARMEKASSVFNVNS 20
   |||||
Db 3 EKKIARMEKASSVFNVNS 21

RESULT 7
PCT-US95-02121-97
; Sequence 97, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Malaria circumsporozoite
; OTHER INFORMATION: 378-398"
PCT-US95-02121-97

Query Match 52.7%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIARMEKASSVFNVNS 20
   |||||
Db 3 EKKIARMEKASSVFNVNS 21

RESULT 8
PCT-US95-13841-20
; Sequence 20, Application PC/TUS9513841
; GENERAL INFORMATION:
```

```
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IgE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-20

Query Match 52.7%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKKIARMEKASSVFNVNS 20
   |||||
Db 3 EKKIARMEKASSVFNVNS 21

RESULT 9
US-08-760-797A-1
; Sequence 1, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
```

```

; TOPOLOGY: linear
US-08-760-797A-3

Query Match          50.9%; Score 86; DB 2; Length 424;
Best Local Similarity 67.7%; Pred. No. 5.5e-05;
Matches 21; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY  2  EKKIAKMEKASSVFNVNNSGSPSLHWSYGLRP 32
      |||| |||| |||| |||| |||| |||| ||||
Db   171  EKKICKMEKCSSVFNVNNS-----SIGLGP 195

RESULT 11
US-08-932-929B-1
; Sequence 1, Application US/08932929B
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBsAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,929B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,797
; FILING DATE: 04-DEC-1996
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-JFWC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-932-929B-1

Query Match          50.9%; Score 86; DB 4; Length 424;
Best Local Similarity 85.7%; Pred. No. 5.5e-05;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  2  EKKIAKMEKASSVFNVNNSGP 22
      |||| |||| |||| |||| |||| ||||
Db   175  EKKICKMEKCSSVFNVNNSRP 195

RESULT 12
US-08-932-929B-3
; Sequence 3, Application US/08932929B
; Patent No. 6169171
; GENERAL INFORMATION:

```

```

; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-313-288B-18

Query Match 48.5%; Score 82; DB 1: Length 412;
Best Local Similarity 78.3%; Pred. No. 0.00021;
Matches 18; Conservative 0; Mismatches 5; Indels

QY 2 EKKIARMEKASSVFNVVNGPSL 24
    ||| |||| ||||| |||
Db 380 EKKICKMEKSSVFNVVNSSIGL 402

RESULT 14
PCT-US95-02121-96
; Sequence 96, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown

```


Job time : 7.98077 secs

;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1..17
;; OTHER INFORMATION: /note= "Malaria circumsporozoite"
;; OTHER INFORMATION: 382-398"
PCT-US95-02121-96

Query Match 46.7%; Score 79; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 K1AKMEKASSVFNVVNS 20
Db 1 K1AKMEKASSVFNVVNS 17

RESULT 15
US-08-817-933A-7
; Sequence 7, Application US/08817933A
; Patent No. 5945104
; GENERAL INFORMATION:
; APPLICANT: STANWORTH, DENIS R.
; APPLICANT: LEWIN, IAN V.
; TITLE OF INVENTION: PEPTIDES FOR ANTI-ALLERGY TREATMENT
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5945104th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,933A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9422294.0
; FILING DATE: 04-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 179-23
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4005
; TELEFAX: (703) 816-4100
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-817-933A-7

Query Match 44.4%; Score 75; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EKK1AKMEKASSVFNV 17
Db 1 EKK1AKMEKASSVFNV 16

Search completed: October 10, 2002, 16:14:03

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 12.8333 Seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-12

Perfect score: 169

Sequence: 1 DEKTIKMEKASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	82	48.5	115	5 Q9TVM7	Q9TVM7 plasmodium
2	82	48.5	115	5 Q9U934	Q9U934 plasmodium
3	82	48.5	115	5 Q25835	Q25835 plasmodium
4	82	48.5	115	5 Q25836	Q25836 plasmodium
5	82	48.5	115	5 Q25837	Q25837 plasmodium
6	82	48.5	115	5 Q25839	Q25839 plasmodium
7	82	48.5	117	5 Q25794	Q25794 plasmodium
8	82	48.5	117	5 Q25795	Q25795 plasmodium
9	82	48.5	117	5 Q25796	Q25796 plasmodium
10	82	48.5	117	5 Q25797	Q25797 plasmodium
11	82	48.5	393	5 Q92255	Q92255 plasmodium
12	82	48.5	406	5 Q25729	Q25729 plasmodium
13	82	48.5	416	5 Q25829	Q25829 plasmodium
14	82	48.5	420	5 Q25838	Q25838 plasmodium
15	82	48.5	420	5 Q25831	Q25831 plasmodium
16	82	48.5	424	5 Q27425	Q27425 plasmodium

17	82	48.5	424	5 Q99256	Q99256 plasmodium
18	82	48.5	432	5 Q27246	Q27246 plasmodium
19	82	48.5	432	5 Q25827	Q25827 plasmodium
20	82	48.5	436	5 Q27325	Q27325 plasmodium
21	82	48.5	436	5 Q25828	Q25828 plasmodium
22	82	48.5	442	5 Q25830	Q25830 plasmodium
23	82	48.5	452	5 Q25834	Q25834 plasmodium
24	81	47.9	80	5 Q9TW97	Q9TW97 plasmodium
25	81	47.9	80	5 Q9TW83	Q9TW83 plasmodium
26	81	47.9	80	5 Q9TW76	Q9TW76 plasmodium
27	81	47.9	80	5 Q9TW01	Q9TW01 plasmodium
28	81	47.9	80	5 Q9TW00	Q9TW00 plasmodium
29	81	47.9	80	5 Q9TVQ0	Q9TVQ0 plasmodium
30	81	47.9	80	5 Q9TVP9	Q9TVP9 plasmodium
31	81	47.9	80	5 Q9TVN9	Q9TVN9 plasmodium
32	81	47.9	80	5 Q9U0Q4	Q9U0Q4 plasmodium
33	81	47.9	80	5 Q9U0Q0	Q9U0Q0 plasmodium
34	81	47.9	80	5 Q9U0P9	Q9U0P9 plasmodium
35	81	47.9	80	5 Q9U0P8	Q9U0P8 plasmodium
36	81	47.9	80	5 Q9U0P7	Q9U0P7 plasmodium
37	81	47.9	80	5 Q9U0P6	Q9U0P6 plasmodium
38	81	47.9	80	5 Q9U0P5	Q9U0P5 plasmodium
39	81	47.9	80	5 Q9U0P4	Q9U0P4 plasmodium
40	81	47.9	80	5 Q9U0P3	Q9U0P3 plasmodium
41	81	47.9	80	5 Q9U0P2	Q9U0P2 plasmodium
42	81	47.9	80	5 Q9U0P1	Q9U0P1 plasmodium
43	78	46.2	80	5 Q9U0Q3	Q9U0Q3 plasmodium
44	78	46.2	80	5 Q9U0Q1	Q9U0Q1 plasmodium
45	77	45.6	79	5 Q9U0Q2	Q9U0Q2 plasmodium

ALIGNMENTS

RESULT 1

Q9TVM7
ID Q9TVM7 PRELIMINARY; PRT; 115 AA.
AC Q9TVM7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=815, 947, AND 808;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwattiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporoite protein of Plasmodium falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83153; AAA29566.1; -;
DR EMBL; M83171; AAA29549.1; -;
DR EMBL; M83151; AAA29564.1; -;
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12925 MW; 697DF63EEBEEDA90 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;

Best Local Similarity 78.3%; Pred. No. 0.00023;

Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIARMEKASSVFNVNSGPSL 24

||||| ||||| ||||| |||||

Db 83 EKKICKMEKCSSVFNVNSSIGL 105

[illegible]

Query Match 48.5%; Score 82; DB 5; Length 117;
Best Local Similarity 78.3%; Pred. No. 0.00023;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNGPSL 24
||||| ||||| ||||| ||||| |||||
Db 85 EKKICKMEKCSSVFNVNNGSGL 107

RESULT 10

Q25797 ID Q25797 PRELIMINARY; PRT; 117 AA.
AC Q25797;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Doolan D.L., Saul A., Good M.F.;
RT "Geographically restricted heterogeneity of the Plasmodium falciparum
circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M77205; AAA29519.2; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
FT NON_TER 1
FT 117
SQ SEQUENCE 117 AA; 13043 MW; 0DA711D86C0B03C1 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 117;
Best Local Similarity 78.3%; Pred. No. 0.00023;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNGPSL 24
||||| ||||| ||||| ||||| |||||
Db 85 EKKICKMEKCSSVFNVNNGSGL 107

RESULT 11

Q99255 ID Q99255 PRELIMINARY; PRT; 393 AA.
AC Q99255;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Lockyer M.J.;
RX MEDLINE=91270295; PubMed=2052038;
RT "Clonal variation in the Plasmodium falciparum circumsporozoite
protein gene.";
RL Mol. Biochem. Parasitol. 45:179-181(1991).
DR EMBL; M57498; AAA63421.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
KW Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 >393 CIRCUMSPOROZOITE PROTEIN.
FT NON_TER 393

SQ SEQUENCE 393 AA; 42263 MW; 45169AE773689037 CRC64;
Query Match 48.5%; Score 82; DB 5; Length 393;
Best Local Similarity 78.3%; Pred. No. 0.00089;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNGPSL 24
||||| ||||| ||||| ||||| |||||
Db 361 EKKICKMEKCSSVFNVNNGSGL 383

RESULT 12

Q25729 ID Q25729 PRELIMINARY; PRT; 408 AA.
AC Q25729;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN.
GN CS.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SANTA LUCIA, SALL1;
RA Qari S.H., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20969; AAA63153.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
KW Malaria.
SQ SEQUENCE 408 AA; 43871 MW; A545BE517822515C CRC64;

Query Match 48.5%; Score 82; DB 5; Length 408;
Best Local Similarity 78.3%; Pred. No. 0.00093;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNGPSL 24
||||| ||||| ||||| ||||| |||||
Db 376 EKKICKMEKCSSVFNVNNGSGL 398

RESULT 13

Q25829 ID Q25829 PRELIMINARY; PRT; 416 AA.
AC Q25829;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAD20;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83172; AAA29550.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
SQ SEQUENCE 416 AA; 44829 MW; D3EF560B2D368DE9 CRC64;

Search completed: October 10, 2002, 16:09:59
Job time : 13.8333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 17.3462 Seconds
(without alignments)
211.311 Million cell updates/sec

Title: US-09-848-834a-12

Perfect score: 169

Sequence: 1 DEKKIAKMEKASSVFVNNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	134.5	79.6	33	15	AAR62715
2	95	56.2	33	16	AAR83570
3	89	52.7	19	22	AAM98951
4	89	52.7	21	10	AAP91504
5	89	52.7	21	16	AAR82586
6	89	52.7	21	16	AAR78920
7	89	52.7	21	16	AAR75955
8	89	52.7	21	16	AAR70912
9	89	52.7	21	17	AAM05612
10	89	52.7	21	18	AAM35440
11	89	52.7	21	20	AAV23252

12	89	52.7	21	21	AAV80071	Pathogen derived T
13	89	52.7	21	21	AAV54553	T helper cell (Th)
14	89	52.7	21	21	AAV58777	Unidentified pepti
15	89	52.7	21	22	AAB99706	Plasmodium falcipa
16	89	52.7	21	22	AAG62428	Plasmodium falcipa
17	89	52.7	21	22	AAG84517	Plasmodium falcipa
18	89	52.7	21	22	AAG8269	Plasmodium falcipa
19	89	52.7	21	22	AAG9366	Plasmodium falcipa
20	89	52.7	21	22	AAB84447	Sequence of T help
21	89	52.7	21	22	AAB98457	Plasmodium falcipa
22	89	52.7	22	16	AAR82077	Malaria CSR3 prote
23	89	52.7	33	22	AAG63663	Peptide comprising
24	89	52.7	33	22	AAG63516	A peptide which ma
25	89	52.7	143	21	AAV49252	N6 polyepitope car
26	89	52.7	218	21	AAV49253	N10 polyepitope ca
27	89	52.7	240	21	AAV49254	N11 polyepitope ca
28	89	52.7	390	21	AAV49255	N19 polyepitope ca
29	86	50.9	424	14	AAR37796	RMS* protein. Syn
30	86	50.9	424	14	AAR37797	RTS* protein. Syn
31	85	50.3	18	21	AAV49259	CD4+ T cell epitop
32	82	48.5	21	15	AAR65375	Helper T cell epit
33	82	48.5	309	12	AAR13175	NS1_81-Rufdeltag.
34	82	48.5	319	11	AAR07945	NS18IRLFAuth plasm
35	82	48.5	319	12	AAR13176	NS1_81-RLfAuth. p
36	82	48.5	327	12	AAR13177	NS1_81-RLfAuth. + (
37	82	48.5	335	12	AAR13178	NS1_81(NVDP)4RLFAu
38	82	48.5	335	12	AAR13179	Sequence encoded b
39	82	48.5	411	9	AAP83144	CS protein of mala
40	82	48.5	412	7	AAP60416	Plasmodium falcipa
41	81	47.9	21	21	AAV70283	Recombinant vaccin
42	81	47.9	350	21	AAV70278	Malaria circumspor
43	79	46.7	17	16	AAR78919	Malaria circumspor
44	79	46.7	17	16	AAV70911	Malaria class II bindi
45	79	46.7	17	21	AAV99032	

ALIGNMENTS

RESULT 1

AAAR62715	
ID	AAAR62715 standard; peptide; 33 AA.
XX	
AC	AAAR62715;
XX	
DT	10-SEP-1995 (first entry)
XX	
DE	LHRH-containing immunogenic peptide.
XX	
KW	Helper T cell epitope; universal immune stimulator; Invasin; haptan;
KW	vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW	androgen-dependent carcinoma; antitumour; infertility;
KW	Plasmodium falciparum circumsporozoite.
XX	
OS	Synthetic.
XX	
FH	Key
Domain	Location/Qualifiers
FT	1..21
FT	/note= "Plasmodium falciparum circumsporozoite"
FT	helper T cell epitope"
FT	Domain
FT	24..33
FT	/note= "LHRH haptan"
XX	
PN	W09425060-A.
XX	
PD	10-NOV-1994.
XX	
PF	28-APR-1994; 94WO-US04832.
XX	
PR	27-APR-1993; 93US-0057166.
PR	14-APR-1994; 94US-0229275.
XX	
PA	(LADD//) LADD A E.

CC least one (I) where this is a major histocompatibility complex (MHC)
 CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
 CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
 CC allergy, graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
 CC or fungal infections; or (ii) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I), e.g. melanoma
 CC and human immunodeficiency virus infection. AAM98898 to AAM99592
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.

XX SQ Sequence 19 AA;
 Query Match 52.7%; Score 89; DB 22; Length 19;
 Best Local Similarity 100.0%; Pred. No. 9.7e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVVNS 20
 Db 1 EKKIAKMEKASSVFNVVNS 19
 |||||

RESULT 4
 AAP91504
 ID AAP91504 standard; peptide; 21 AA.
 XX AC AAP91504;
 XX DT 13-MAR-1992 (first entry)
 XX DE Sequence of modified Plasmodium circumsporozoite T-cell epitope CS.T3.
 XX DE Circumsporozoite peptide; T-cell epitope; immunogenic composition;
 XX KW vaccine.
 XX OS Plasmodium falciparum.
 XX PH Key Location/Qualifiers
 FT Misc-difference 1..2 /note= "May be H-Asp-Ile, H-Ile, or H-"
 FT Misc-difference 19..21 /note= "May be Val-Asn-Ser-OH, Val-Asn-OH,
 FT Val-OH or -OH"
 XX EP343460-A.
 XX 29-NOV-1989.
 XX 12-MAY-1989; 89EP-0108618.
 XX 24-MAY-1988; 88GB-0012214.
 XX (HOFF) HOFFMANN-LA ROCHE AG.
 XX Sinigaglia F;
 XX WPI; 1989-349561/48.
 XX Modified Plasmodium CS peptide - used as a universally recognised
 PT T-cell epitope in vaccines to elicit an immune response against
 PT pathogenic agents
 XX Claim 1; Page 16; 23pp; English.
 XX Peptide CS.T3 (residues 1-21) corresponds to residues 378-398 of the
 CC CS protein from P. falciparum but contains 2 Ala residues in place

CC of the native protein's Cys residues at positions 384 and 389. Also
 CC claimed is AAP91504 (or modified forms, see FT) associated with an
 CC antigenic structure representing a B-cell epitope, pref. a multiple
 CC antigenic peptide, esp. multimers of the repeat sequences NAMP
 CC present in P. falciparum CS protein.

XX SQ Sequence 21 AA;
 Query Match 52.7%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVVNS 20
 Db 3 EKKIAKMEKASSVFNVVNS 21
 |||||

RESULT 5
 AAR82586
 ID AAR82586 standard; peptide; 21 AA.
 XX AC AAR82586;
 XX DT 13-JUN-1996 (first entry)
 XX DE Plasmodium falciparum circumsporozoite helper T cell epitope, PF.
 XX DE IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 XX KW vaccine; allergy; antibody; constant heavy chain.
 XX OS Plasmodium falciparum.
 XX PN WO9526365-A1.
 XX PD 05-OCT-1995.
 XX 24-MAR-1995; 95WO-US03741.
 XX 25-OCT-1994; 94US-0328912.
 XX 28-MAR-1994; 94US-0218461.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY;
 XX WPI; 1995-351297/45.
 XX Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper
 PT T cell epitope - useful for eliciting antibody prodn. for allergy
 PT treatment
 XX Claim 3; Page 23; 87pp; English.

XX AAR82571-91 are helper T cell epitopes which can be used in the
 CC preparation of a peptide immunogen that is useful in vaccines for
 CC treating allergic reactions. In the immunogen an IgE CH4 peptide
 CC is attached C-terminally to a series of amino acids including a
 CC helper T cell epitope. The immunogen may also opt. contain a fatty
 CC acid or fatty acid derivative, an invasion domain or alpha-NH₂. The
 CC immunogen produces high titres of antibodies to the effector site
 CC in human IgE heavy chain (the CH4 domain peptide) which inhibit mast
 CC cell activation and reduce allergen-induced IgE prodn. The immunogens
 CC may be used in either a radially branching multimeric form or a
 CC linearly arranged monomeric form.

XX SQ Sequence 21 AA;
 Query Match 52.7%; Score 89; DB 16; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVVNS 20
 |||||

```

Dh      3 EKKIAKMEKASSVFNVNS 21

RESULT 6
AAR78920
ID      AAR78920 standard; peptide; 21 AA.
XX
AC      AAR78920;
XX
DT      27-MAR-1996 (first entry)
XX
DE      Malaria circumsporozoite 378-398 helper T lymphocyte epitope.
XX
KW      378-398; cytotoxic T; CTL; epitope; helper T; HTL; cell;
KW      lymphocyte; antigens; treatment; disease prevention; tumours;
KW      viruses; parasites; malaria circumsporozoite; hepatitis B.
XX
OS      Malaria circumsporozoite.
XX
PN      WO9522317-A1.
XX
PD      24-AUG-1995.
XX
PF      16-FEB-1995; 95WO-US02121.
XX
PR      16-FEB-1994; 94US-0197484.
XX
PA      (CYTE-) CYTEL CORP.
XX
PI      Ceut RW, Grey H, Sette AD, Vitiello MA;
XX
DR      WPI; 1995-302545/39.
XX
PT      Compsn. inducing cytotoxic T lymphocyte response to pref. viral,
PT      bacterial, parasitic or tumour antigens - useful in the treatment
PT      and prevention of diseases associated with the antigen e.g.
PT      hepatitis B
XX
PS      Disclosure; Page 29; 109pp; English.
XX
CC      A compsn. which induces a cytotoxic T lymphocyte (CTL) response to
CC      an antigen (Ag) in a mammal comprises, a CTL Ag response inducing
CC      peptide and a lipid conjugated helper T cell inducing peptide (i.e.
CC      AAR78918-R78922). The compsn. induces a CTL response to bacterial,
CC      viral or tumour Ags, and is therefore useful in the treatment and
CC      prevention of diseases associated with the Ag, e.g. hepatitis B.
XX
SQ      Sequence 21 AA;

Query Match 52.7%; Score 89; DB 16; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAKMEKASSVFNVNS 20
        |||||
Dh      3 EKKIAKMEKASSVFNVNS 21

RESULT 7
AAR75955
ID      AAR75955 standard; Peptide; 21 AA.
XX
AC      AAR75955;
XX
DT      06-MAR-1996 (first entry)
XX
DE      P. falciparum CS protein epitope residues 378-398.
XX
KW      MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;
KW      cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.
XX
OS      Plasmodium falciparum.
XX

Query Match 52.7%; Score 89; DB 16; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAKMEKASSVFNVNS 20
        |||||
Dh      3 EKKIAKMEKASSVFNVNS 21

RESULT 8
AAR70912
ID      AAR70912 standard; Protein; 21 AA.
XX
AC      AAR70912;
XX
DT      09-OCT-1995 (first entry)
XX
DE      Malaria circumsporozoite 378-398 T helper peptide.
XX
KW      Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW      malaria circumsporozoite 378-398; T helper peptide.
XX
OS      Malaria circumsporozoite.
XX
PN      WO9504542-A.
XX
PD      16-FEB-1995.
XX
PF      02-AUG-1994; 94WO-US08721.
XX
PR      06-AUG-1993; 93US-0103623.
XX
PA      (CYTE-) CYTEL CORP.
XX
PI      Fikes JD, Livingston BD, Sette AD, Sidney JC;
XX
DR      WPI; 1995-090681/12.
XX
PT      Human melanoma antigen, MAGE-1, peptide(s) - useful for
PT      stimulating immune response against melanoma
XX
PS      Disclosure; Page 13; 59pp; English.

```

XX The T helper peptides described in AAR70910-R70914 are used in
CC conjunction with the C-terminal MAGE-1 peptides described in
CC AAR70915 to AAR70969. Compsns. containing the T helper and MAGE-1
CC peptides can be administered, as a vaccine to patients susceptible
CC to MAGE associated tumours, e.g. melanomas.
XX
SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 16; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNS 20
| | | | | | | | | | | | | | | | | | | | |
Db 3 EKKIAMEKASSVFNVNS 21

RESULT 9
AAW05612
ID AAW05612 standard; peptide; 21 AA.
XX
XX AAW05612;
XX
XX 10-DEC-1996 (first entry)
XX
XX Circumsporozoite helper T cell epitope.
XX
XX Immunoglobulin; IgE; membrane protein; human; epsilon chain; hepatitis B;
KW membrane anchoring domain; helper T cell; surface antigen; core antigen;
KW pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;
KW Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;
KW diphtheria toxin; plasmodium falciparum; circumsporozoite; E. coli f7af;
KW schistosoma mansoni; triose phosphate isomerase; allergenic reaction;
KW allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;
KW antihistamine; decongestant; beta-2 agonist; immunosuppression;
KW corticosteroid.
XX
XX Plasmodium falciparum.
OS
XX
XX W09612740-A1.
XX
XX 02-MAY-1996.
XX
XX 25-OCT-1995; 95WO-US13841.
XX
XX 25-OCT-1994; 94US-0328519.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Walfield AM, Wang CY;
XX
XX WPI; 1996-230555/23.
XX
XX Peptide immunogen useful in treatment of allergy - comprises
PT membrane-bound IgE epsilon-chain peptide synthesised linearly in
PT tandem with T helper epitope peptide
XX
XX Claim 2; Page 19; 53pp; English.

AAW05957-W05616 represent helper T cell epitopes used in the peptide
CC immunogens of the invention. This sequence represents the Plasmodium
CC falciparum circumsporozoite helper T cell antigen. The peptides of the
CC invention contain one of these sequences, and a membrane-bound
CC immunoglobulin E (IgE) fragment (see AAW05595 and AAW05596). The
CC peptide immunogens of the invention can be used in vaccines for the
CC immunotherapeutic treatment of allergic reactions, including allergic
CC rhinitis, food allergies, anaphylaxis, or virally-induced asthma. The
CC immunogens overcome the short effective period of antihistamines,
CC decongestants, and beta-2 agonists, while preventing the broad
CC immunosuppression of corticosteroids. The peptides do not have the
CC potential side effects of restlessness or sedation (associated with
CC antihistamines), associated increased morbidity in asthmatics (as seen

CC with beta-2 agonists) and adverse hormonal activities (observed in
CC corticosteroid users).
XX
SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNS 20
| | | | | | | | | | | | | | | | | | | | |
Db 3 EKKIAMEKASSVFNVNS 21

RESULT 10
AAW35440
ID AAW35440 standard; peptide; 21 AA.
XX
XX AAW35440;
XX
XX 22-APR-1998 (first entry)
XX
XX T-cell stimulatory peptide from Plasmodium falciparum.
XX
XX T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
KW scaffold; inhibition; metastasis; wound healing; solid phase.
XX
XX Plasmodium falciparum.
XX
XX W09738011-A1.
XX
XX 16-OCT-1997.
XX
XX 03-APR-1997; 97WO-DE00146.
XX
XX 03-APR-1996; 96DK-0000398.
XX
XX (PEPR-) PEPRESEARCH AS.
XX
XX Heegaard PMH, Jakobsen PH;
XX
XX WPI; 1997-512645/47.
XX
XX Non-dendritic peptide carrier linked to a solid phase - useful as a
PT diagnostic agent and as a scaffold for production of chemical
PT derivatives
XX
XX Claim 30; Page 199; 262pp; English.

A non-dendritic peptide carrier (A) has been developed which is coupled
CC through a linker to a solid phase, forming a complex of (A)-solid phase.
CC Where (A) comprises 10-50 amino acids capable of forming a secondary
CC structure in a benign buffer after liberation from the solid phase, and
CC further the (A)-solid phase complex comprises an immunogenic substance
CC and/or an immune mediator coupled on (A). The present sequence
CC represents a specifically claimed T-cell stimulatory peptide from the
CC present invention. An (A)-solid phase complex can be used as a scaffold
CC for the production of chemical derivatives, characterised by covalently
CC attaching molecules at attachment points. Alternatively (A) is used as
CC a scaffold-peptide for the incorporation into an immunostimulating
CC complex (Iscom) resulting in an (A)-Iscom complex which is used for the
CC chemical coupling of antigenic substances in an aqueous solution by
CC conjugation. (A) derivatised with one or more peptides having
CC fibronectin-, laminin- or vitronectin-like binding activities can be
CC used for the promotion of cell-attachment to plastic surfaces, in
CC particular to inhibit tumour growth and metastasis, and for promotion
CC of wound healing. Also a derivatised (A) can be used for the selection
CC of specifically-binding aptamers or as a diagnostic agent. Such
CC diagnostic-(A) molecules could be used to detect molecules derived from
CC or indicative of pregnancy or of a disease, such as an infectious,
CC autoimmune or cancerous disease.
XX
XX Sequence 21 AA;

XX

DT	15-MAY-2000	(first entry)
XX		
DE	Pathogen derived Th epitope SEQ ID NO:78.	
XX		
KW	Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;	
KW	immunogenic; immunostimulatory; carrier protein; helper T cell epitope;	
KW	antibody; allergy; allergic disease; immunisation; anti-allergic;	
KW	anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.	
XX		
OS	Unidentified.	
XX		
PN	WO9967293-A1.	
XX		
PD	29-DEC-1999.	
XX		
PF	21-JUN-1999; 99WO-US13959.	
XX		
PR	20-JUN-1998; 98US-0100287.	
XX		
PA	(UNBI-) UNITED BIOMEDICAL INC.	
XX		
PI	Wang CY, Walfield AM;	
XX		
DR	WPT; 2000-160578/14.	
XX		
PT	New antigenic peptide from the CH3 domain of immunoglobulin E, fusions	
PT	for immunization against allergy	
XX		
PS	Claim 11; Page 79; 155pp; English.	
XX		
CC	The present invention describes immunoglobulin E (IgE)-CH3 domain	
CC	antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic	
CC	and anti-asthmatic properties. (I) induces polyclonal antibodies	
CC	specific for a target effector site on the epsilon-heavy chain of IgE,	
CC	and so preventing triggering and activation of mast cells and basophils	
CC	and downregulation of IgE synthesis. Conjugates, or fusion peptides,	
CC	containing (I) are used for active immunisation against IgE-mediated	
CC	allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy	
CC	dermatitis. Nucleic acids that encode these compounds are useful for	
CC	recombinant production of corresponding peptides or in DNA vaccines.	
CC	Conjugates of (I) that include a promiscuous T helper cell epitope	
CC	(functional in genetically diverse subjects), in addition to a B cell	
CC	target epitope, have increased immunogenicity and may include cyclic	
CC	constraints (disulfide bridge) to stabilise conformational features and	
CC	maximize cross-reactivity to the natural target. They induce safe	
CC	(non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino	
CC	acid sequences used in the exemplification of the present invention.	
XX		
SQ	Sequence 21 AA;	
	Query Match 52.7%; Score 89; DB 21; Length 21;	
	Best Local Similarity 100.0%; Pred. No. 1.le-06;	
	Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	2 EKTIKMEKASSVFNVNS 20	
DB	3 EKTIKMEKASSVFNVNS 21	
RESULT 13		
AAY54553		
ID	AAY54553 standard; peptide; 21 AA.	
XX		
AC	AAY54553;	
XX		
DT	25-APR-2000 (first entry)	
XX		
DE	T helper cell (Th) epitope of Plasmodium flaciparum cicumsporozoite.	
XX		
KW	Human; CD4 protein; antigenic peptide; CDR2-like domain; apoptosis;	
KW	synchronia formation; human immune deficiency virus; HIV binding;	
KW	CD4-Class II interaction; immunisation; CD4 surface complex;	
KW	immune response; transplant rejection; autoimmune disease; psoriasis;	

KW rheumatoid arthritis; systemic lupus erythematosus; ciumsporozoite.
 XX Plasmodium falciparum.
 OS WO9967294-A1.
 XX 29-DEC-1999.
 XX 21-JUN-1999; 99WO-US14030.
 XX 20-JUN-1998; 98US-0100409.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY;
 PI WPI; 2000-160579/14.
 DR New antigenic peptide from the CDR2 domain of CD4, for immunization
 XX against e.g. human immune deficiency virus -
 PT Claim 11; Page 65; 106pp; English.
 XX The present sequence represents a broadly reactive promiscuous T helper
 CC cell (Th) epitope derived from Plasmodium falciparum circumsporozoite. It
 CC is conjugated to antigenic peptides derived from the CDR2-like domain of
 CC the human CD4 protein. These antigenic peptides present neutralising
 CC receptor/co-receptor effector sites of the CDR2-like domain. The
 CC peptides evoke effective antibody responses by having optimised
 CC site-specificity. The induced antibodies block human immune deficiency
 CC virus (HIV) binding and syncytia formation. They may also block
 CC CD4-Class II interactions with other cells, deliver signals to T
 CC cells (inhibiting normal CD4+-mediated immunoregulatory functions) or
 CC induce apoptosis of CD4 cells by simultaneous engagement of T cell
 CC receptors. Conjugates and peptides containing the antigenic peptides are
 CC used for active immunisation to generate antibodies against CD4 surface
 CC complexes, especially to prevent binding of HIV to CD4 and thus HIV
 CC infection, but also to treat undesirable immune responses such as
 CC transplant rejection, or autoimmune diseases (rheumatoid arthritis,
 CC systemic lupus erythematosus or psoriasis). These conjugates produce
 CC high-titre antibodies which are broadly neutralising against primary
 CC isolates from all classes of HIV-1 and of HIV-2.
 XX Sequence 21 AA;
 SQ Query Match 52.7%; Score 89; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 EKKIAMEKASSVFNVVNS 20
 Db 3 EKKIAMEKASSVFNVVNS 21
 RESULT 14
 AAY58777 standard; Peptide; 21 AA.
 ID AAY58777
 XX AAY58777;
 AC AAY58777;
 XX 25-APR-2000 (first entry)
 DT Unidentified peptide.
 XX Helper T cell; Th epitope; feed additive; growth promotion;
 KW somatostatin.
 XX Unidentified.
 OS WO9966950-A1.
 XX 29-DEC-1999.
 PD
 XX

PF 21-JUN-1999; 99WO-US13923.
 XX 20-JUN-1998; 98US-0100415.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY;
 PI WPI; 2000-160560/14.
 DR New somatostatin helper T-cell epitope conjugate for raising
 XX anti-somatostatin antibodies to enhance growth rate in animal by
 PT reducing growth inhibitory activity of somatostatin -
 XX Disclosure; Page 53; 59pp; English.
 PS The present sequence is that of an unidentified peptide of the
 XX invention. The invention relates to peptide compositions (see
 CC AAY58739-66) useful as immunogens for growth promotion in farm
 CC animals. The immunogenic peptides contain helper T cell epitopes
 CC which comprise multiple class II MHC motifs and have somatostatin
 CC at either the C- or N-terminus. They may also include an invasion
 CC domain which acts as a general immune stimulator. The helper T
 CC cell epitopes and the invasion domain enhance the immune response
 CC against the somatostatin self-peptide.
 XX Sequence 21 AA;
 SQ Query Match 52.7%; Score 89; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 EKKIAMEKASSVFNVVNS 20
 Db 3 EKKIAMEKASSVFNVVNS 21
 RESULT 15
 AAB99706
 ID AAB99706 standard; peptide; 21 AA.
 XX AAB99706;
 AC AAB99706;
 XX 06-SEP-2001 (first entry)
 DT Plasmodium falciparum CS protein fragment 378-398 SEQ ID NO:27.
 DE Human leukocyte antigen A2 binding peptide; HLA class I A2; CTL;
 XX cytotoxic T-cell lymphocyte; tumour associated antigen; CEA; HER2/neu;
 KW MAGE2; MAGE3; p53; vaccine; cancer; cytostatic; immunomodulator;
 KW immunotherapy; immune response.
 XX Plasmodium falciparum.
 OS WO200141741-A1.
 XX 14-JUN-2001.
 PD 13-DEC-2000; 2000WO-US34318.
 XX 13-DEC-1999; 99US-0170448.
 PR 05-APR-2000; 2000US-0543608.
 PR 30-MAY-2000; 2000US-0583200.
 XX (EPIM-) EPIMMUNE INC.
 PA Fikes J, Sette A, Sidney J, Southwood S, Celis E, Keogh E;
 XX Chesnut R;
 PI WPI; 2001-381489/40.
 DR Compositions for use in a vaccine for treating, e.g., breast, lung and
 PT colon cancer comprises at least one peptide that comprises an isolated

PT epitope of a tumor-associated antigen -

XX
PS
XX Disclosure; Page 31; 86pp; English.

XX
CC The present invention describes a composition (I) comprising at least
CC one peptide that comprises an isolated, prepared epitope consisting of
CC a sequence selected from 25 short amino acid sequences given in AAB99680
CC to AAB99704. Also described are: (I) a composition (II) comprising one
CC or more peptides, and further comprising at least two epitopes selected
CC from the 25 short amino acid sequences (as above), where each of the one
CC or more peptides comprise less than 50 contiguous amino acids that have
CC 100% identity with a native peptide sequence; and (2) a vaccine
CC composition (III) comprising an epitope selected from the 25 short amino
CC acid sequences (as above) and a pharmaceutical excipient. (I) has
CC cytostatic and immunomodulatory activities and can be used in vaccine
CC production and immunotherapy. The peptide epitope compositions (I)-(II)
CC are useful for monitoring an immune response to a tumour associated
CC antigen or when one or more peptides are combined to create a vaccine
CC (III) that stimulates the cellular arm of the immune system. In
CC particular, the vaccine mediates immune responses against tumours in
CC individuals who bear an allele of the human leukocyte antigen (HLA)-A2
CC supertype and improve the standard of care for patients being treated
CC for breast, colon, or lung cancer. The present sequence represents a
CC plasmodium falciparum CS protein fragment which is given in the
CC exemplification of the present invention.

XX
SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVYNS 20
| | | | | | | | | | | | | | | | | | | | | |
Db 3 EKKIAKMEKASSVFNVYNS 21

Search completed: October 10, 2002, 16:05:11
Job time : 19.3462 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 4.30128 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-12

Perfect score: 169
Sequence: 1 DEKTKAKMEKASSVFNVNVSGPSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	48.5	388	1	CSP_PLARE
2	82	48.5	397	1	CSP_PLAFO
3	82	48.5	412	1	CSP_PLAFA
4	82	48.5	424	1	CSP_PLAFT
5	82	48.5	442	1	CSP_PLAFL
6	56.5	33.4	92	1	GONI_RAT
7	55	32.5	339	1	CSP_PLABE
8	55	32.5	347	1	CSP_PLABA
9	55	32.5	367	1	CSP_PLAYO
10	54	32.0	67	1	GONI_MACMU
11	54	32.0	90	1	GONI_MOUSE
12	54	32.0	91	1	GONI_PIG
13	54	32.0	92	1	GONI_HUMAN
14	54	32.0	95	1	GONI_PAGMA
15	52	30.8	61	1	GONI_SHEEP
16	52	30.8	63	1	GONI_MESAU
17	52	30.8	89	1	GONI_XENLA
18	52	30.8	92	1	GONI_TUPGB
19	52	30.8	994	1	DPO2_KIULA
20	51	30.2	92	1	GONI_CAVPO
21	51	30.2	393	1	CSP_PLABR
22	51	30.2	429	1	CSP_PLAMA
23	50	29.6	95	1	GONI_SPAU
24	50	29.6	2339	1	RPC1_PLAFA
25	49.5	29.3	630	1	ESAB_TRYBB
26	49.5	29.3	3176	1	CA36_HUMAN
27	49	29.0	95	1	GONI_MORSA
28	49	29.0	532	1	INVA_YEAST
29	49	29.0	1146	1	ITAL_DROME
30	48.5	28.7	630	1	ESAB_TRYEQ
31	48	28.4	10	1	GONI_ALLMI
32	48	28.4	92	1	GONI_CHICK
33	48	28.4	99	1	GONI_DICLA

34	48	28.4	3133	1	HMCT_BOMMO
35	47.5	28.1	94	1	GONI_HAPBU
36	47.5	28.1	668	1	VNCS_MEVA
37	47.5	28.1	725	1	ADDB_MOUSE
38	47.5	28.1	725	1	ADDB_RAT
39	47	27.8	275	1	APAH_HAEIN
40	47	27.8	1032	1	ITA4_XENLA
41	46.5	27.5	110	1	YHBJ_ACTAC
42	46.5	27.5	621	1	ULPL_YEAST
43	46.5	27.5	1392	1	LYS2_YEAST
44	46	27.2	212	1	ATPO_YEAST
45	46	27.2	218	1	UNG_SFVKA

ALIGNMENTS

RESULT 1					
CSP_PLARE					
ID	CSP_PLARE	STANDARD;	PRT;	388	AA.
AC	P26694;				
DT	01-AUG-1992 (Rel. 23, Created)				
DT	01-AUG-1992 (Rel. 23, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Circumsporozoite protein precursor (CS).				
OS	Plasmodium reichenowi.				
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				
OX	NCBI_TaxID=5854;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=91201303; PubMed=2016283;				
RA	Lal A.A., Goldman I.F.;				
RT	"Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria parasite evolutionarily related to the human malaria parasite Plasmodium falciparum.";				
RL	J. Biol. Chem. 266:6686-6689(1991).				
CC	-!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).				
CC	-!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.				
CC	-!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; M60972; AAA29561.1; -				
DR	PIR; A39756; A39756; Crmsprzoite.				
DR	InterPro; IPR003087; Crmsprzoite.				
DR	InterPro; IPR000884; TSPL.				
DR	Pfam; PF00090; tsp_1; 1.				
DR	PRINTS; PR01303; CRMSPRZOITE.				
DR	SMART; SM00209; TSPL; 1.				
KW	Malaria; Sporozoite; Repeat; Signal.				
FT	SIGNAL 1 16 PROBABLE.				
FT	CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.				
FT	DOMAIN 120 267 62 X 4 AA TANDEM REPEATS OF N-A-N-P.				
SQ	SEQUENCE 388 AA; 42245 MW; C03EEFBE2E35604 CRC64;				

Query Match					
Best Local Similarity 48.5%; Score 82; DB 1; Length 388;					
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;					
QY	2	EEKIAKMEKASSVFNVNVSGPSL	24		
DB	356	EKKICKMEKCSSVFNVNVSSIGL	378		

Query Match 48.5%; Score 82; DB 1; Length 397;
 Best Local Similarity 78.3%; Pred. No. 0.00037;
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNSGPSL 24
 ||||| ||||| ||||| |||||
 DB 365 EKKICKMEKCSSVFNVNSSIGL 387

RESULT 3

CSP_PLAFA STANDARD; PRT; 412 AA.
 AC P02893;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84250215; PubMed=6204383;
 RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
 RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
 RA Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
 RT "Structure of the gene encoding the immunodominant surface antigen on
 the sporozoite of the human malaria parasite Plasmodium falciparum.";
 RL Science 225:593-599(1984).
 CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 CC VERTEBRATE HOST).
 CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K02194; AAA29524.1; -.
 CC PIR; A03388; OZZQAF.
 CC DR InterPro: IPR003067; Crcmsprzoite.
 CC DR InterPro: IPR000884; TSP1.
 CC DR Pfam: PF00090; tsp.1; 1.
 CC DR PRINTS: PR01303; CRCMSPRZOITE.
 CC DR SMART: SM00209; TSP1; 1.
 CC Malaria; Sporozoite; Repeat; Signal.
 CC FT SIGNAL 1 16
 CC CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
 CC FT DOMAIN 123 290 41 X 4 AA TANDEM REPEATS OF P-N-A-N.
 CC SEQUENCE 412 AA; 44420 MW; 1EEED3DE90965F8 CRC64;

Query Match 48.5%; Score 82; DB 1; Length 412;
 Best Local Similarity 78.3%; Pred. No. 0.00039;
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNSGPSL 24
 ||||| ||||| ||||| |||||
 DB 380 EKKICKMEKCSSVFNVNSSIGL 402

RESULT 4

CSP_PLAFA STANDARD; PRT; 424 AA.
 ID CSP_PLAFA
 AC P13814;

RESULT 2
 CSP_PLAFA STANDARD; PRT; 397 AA.
 AC P19597; Q25798;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium falciparum (isolate NF54).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5843;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89345189; PubMed=2668895;
 RA Campbell J.R.;
 RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria
 candidate vaccine antigen.";
 RL Nucleic Acids Res. 17:5854-5854(1989).
 RN [2]
 RP REVISIONS.
 RA Campbell J.R.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92155298; PubMed=1346766;
 RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
 RA Thomas A.W., Baqar S., Cochran M.A., Thanassi J., Levine M.M.,
 RA Hackett C.S.;
 RT "Plasmodium falciparum: in vitro characterization and human
 infectivity of a cloned line.";
 RL Exp. Parasitol. 74:159-168(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89364998; PubMed=2671723;
 RA Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;
 RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
 isolate used in malaria vaccine trials.";
 RL Mol. Biochem. Parasitol. 35:195-190(1989).
 CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
 CC SURFACE ANTIGEN ON THE SPOOROZOITE (THE INFECTIVE STAGE OF THE
 CC MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
 CC VERTEBRATE HOST).
 CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
 CC ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
 CC WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
 CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X15363; CAA33421.1; -.
 CC DR EMBL; M83886; AAA29521.1; -.
 CC DR EMBL; M22982; AAA29527.1; -.
 CC DR PIR; S05428; S05428.
 CC DR PIR; A45527; A45527.
 CC DR InterPro: IPR003067; Crcmsprzoite.
 CC DR InterPro: IPR000884; TSP1.
 CC DR Pfam: PF00090; tsp.1; 1.
 CC DR PRINTS: PR01303; CRCMSPRZOITE.
 CC DR SMART: SM00209; TSP1; 1.
 CC Malaria; Sporozoite; Repeat; Signal.
 CC FT SIGNAL 1 16 PROBABLE.
 CC CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
 CC FT DOMAIN 101 272 43 X 4 AA TANDEM REPEATS OF N-A-N-P.
 CC CONFLICT 194 194 A -> ANPNANPNA (IN REF. 4).
 CC SEQUENCE *397 AA; 42646 MW; 9E81146F59BCEA3 CRC64;

QY 2 EKKIAMEKASSVFNVNSGPSL 24
 ||||| ||||| ||||| |||||
 DB 380 EKKICKMEKCSSVFNVNSSIGL 402

KW Signal. 1 1
 FT NON_TER <1 5
 FT SIGNAL 6
 FT CHAIN 6
 FT CHAIN 6
 FT PEPTIDE 15
 FT PEPTIDE 19
 FT ACT_SITE 8
 FT MOD_RES 6
 FT MOD_RES 15
 FT NON_TER 67
 FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
 Query Match 32.0%; Score 54; DB 1; Length 67;
 Best Local Similarity 90.0%; Pred. No. 0.54;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 23 SLHWSYGLRP 32
 Db 5 SQHWSYGLRP 14
 RESULT 11
 GONL_MOUSE STANDARD; PRT; 90 AA.
 AC P13562;
 DT 01-JAN-1990 (Rel. 13; Created)
 DT 01-JAN-1990 (Rel. 13; Last sequence update)
 DT 30-MAY-2000 (Rel. 39; Last annotation update)
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
 DE I].
 GN GNRH1 OR GNRH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87069528; PubMed=3024317;
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
 RA Phillips H.S., Nikolics K., Seeburg P.H.;
 RT "A deletion truncating the gonadotropin-releasing hormone gene is
 RT responsible for hypogonadism in the hp9 mouse."
 RL Science 234:1366-1371(1986).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M14872; AAA37717.1; -.
 CC DR MGD; MGI:95789; GnRH.
 DR InterPro: IPR002012; GnRH.
 DR InterPro: IPR004079; GonadolibereinI.
 DR Pfam: PF00446; GnRH; 1.
 DR PRINTS: PR01541; GONADOLIBERNI.
 DR PROSITE: PS00473; GNRH; 1.
 DR KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 90 PROGONADOLIBERIN I.

FT PEPTIDE 22 31 GONADOLIBERIN I.
 FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
 FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;
 Query Match 32.0%; Score 54; DB 1; Length 90;
 Best Local Similarity 90.0%; Pred. No. 0.74;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 23 SLHWSYGLRP 32
 Db 21 SQHWSYGLRP 30
 RESULT 12
 GONL_PIG STANDARD; PRT; 91 AA.
 AC P49921;
 DT 01-OCT-1996 (Rel. 34; Created)
 DT 01-OCT-1996 (Rel. 34; Last sequence update)
 DT 30-MAY-2000 (Rel. 39; Last annotation update)
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
 GN GNRH1 OR GNRH.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RA Weesner G.D., Matteri R.L., Becker B.A.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=72114303; PubMed=4946067;
 RA Baba Y., Matsuo H., Schally A.V.;
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.
 RT Confirmation of the proposed structure by conventional sequential
 RT analyses."
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
 RN [3]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72065376; PubMed=4942726;
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 RT phase method."
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
 RN [4]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72117544; PubMed=4946275;
 RA Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone."
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L32864; AAA31066.1; -.

DR PIR: A01411; RHPGS.
 DR InterPro: IPR002012; GnRH.
 DR Pfam: PF00446; GnRH; 1.
 DR PRINTS: PRO1541; GONADOLIBRN1.
 DR PROSITE: PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 91 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

Query Match 32.0%; Score 54; DB 1; Length 91;
 Best Local Similarity 90.0%; Pred. No. 0.75;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32
 I I I I I I I I
 Db 23 SQHWSYGLRP 32

RESULT 13
 GONL_HUMAN
 ID GONL_HUMAN STANDARD; PRT; 92 AA.
 AC P01148;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Progadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
 peptide I].
 DE GNRH1 OR GNRH OR LHRH.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89366682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 hormone gene."
 RL Nucleic Acids Res. 17:6403-6403(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 of gonadotropin-releasing hormone and prolactin release-inhibiting
 factor in human and rat."
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85012739; PubMed=6090951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 releasing hormone."
 RL Nature 311:666-668(1984).
 RN [4]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=83126573; PubMed=6760865;
 RA Tan L., Rousseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide
 biosynthesized in the human placenta."
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES

CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 CC Lutrepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm
 CC (Serono).
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X01059; CAA25526.1; -.
 DR EMBL: M12578; AAA35916.1; -.
 DR EMBL: X15215; CAA33285.1; -.
 DR PIR: A01410; RHUG.
 DR PIR: A26173; A26173.
 DR PIR: S05308; S05308.
 DR MIM: 152760; -.
 DR InterPro: IPR002012; GnRH.
 DR InterPro: IPR004079; Gonadoliberin1.
 DR Pfam: PF00446; GnRH; 1.
 DR PRINTS: PRO1541; GONADOLIBRN1.
 DR PROSITE: PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pharmaceutical; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 92 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 FT CONFLICT 16 16 W -> S (IN REF. 3).
 SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 32.0%; Score 54; DB 1; Length 92;
 Best Local Similarity 90.0%; Pred. No. 0.75;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32
 I I I I I I I I
 Db 23 SQHWSYGLRP 32

RESULT 14
 GONL_PAGMA
 ID GONL_PAGMA STANDARD; PRT; 95 AA.
 AC P70074;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
 (LH-RH I) (Luliberin I).
 DE GNRH1.
 OS Pagrus major (Red sea bream) (Chrysophrys major).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
 OC Sparidae; Pagrus.
 OX NCBI_TaxID=143350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Okuzawa K., Granneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D86582; BAA13129.1; -.
CC InterPro: IPR002012; GNRH.
CC Pfam: PF00446; GNRH; 1.
CC PROSITE: PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 24 95
FT PEPTIDE 24 33
FT PEPTIDE 37 95
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 33 33
FT SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;
Query Match 32.0%; Score 54; DB 1; Length 95;
Best Local Similarity 52.6%; Pred. No. 0.78; 7; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 7;
QY 14 VENVVNSGFSLHWSYGLRP 32
Db 14 VMVMSQCCQHWISGLSP 32
RESULT 15
ID GON1_SHEEP STANDARD; PRT; 61 AA.
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor (Contains: Gonadoliberein I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I]
DE (Fragment).
DE GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN-WESTERN RANGE; TISSUE-Hypothalamus;
RA Rodriguez R.E.; Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-10.
RX MEDLINE=72094314; PubMed=4550508;
RA Burqus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vale W., Guillemain R.;
RT "Primary structure of the ovine hypothalamic luteinizing hormone-
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-Edman degradation).";
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U02517; AAA03433.1; -.
CC PIR: A93780; RUSHG
CC InterPro: IPR002012; GNRH.
CC Pfam: PF00446; GNRH; 1.
CC PROSITE: PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >61
FT PEPTIDE 1 10
FT PEPTIDE 14 >61
FT ACT_SITE 3 3
FT ACT_SITE 3 3
FT MOD_RES 1 1
FT MOD_RES 10 10
FT NON_TER 61 61
FT SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;
Query Match 30.8%; Score 52; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 HWSYGLRP 32
Db 2 HWSYGLRP 9

Search completed: October 10, 2002, 16:06:31
Job time : 5.30128 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 7.6859 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKTKAKMEKASVFNVNVGSLHWSYGLRPX 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	48.5	388	2 A39756	circumsporozoite p
2	82	48.5	405	2 S05428	circumsporozoite p
3	82	48.5	412	1 OZZQAF	circumsporozoite p
4	82	48.5	424	2 A54533	circumsporozoite p
5	82	48.5	442	2 A54529	circumsporozoite p
6	56.5	33.4	92	1 RHRTG	gonadoliberin prec
7	55	32.5	264	2 A4969	circumsporozoite p
8	55	32.5	332	1 OZZQMB	circumsporozoite p
9	55	32.5	348	1 OZZQBB	circumsporozoite p
10	55	32.5	360	2 AD2528	circumsporozoite p
11	55	32.5	367	1 OZZQMY	integrase-recombin
12	54	32.0	67	2 I78541	circumsporozoite p
13	54	32.0	90	1 RHMSG	gonadoliberin prec
14	54	32.0	92	1 RHSHG	gonadoliberin prec
15	52	30.8	10	1 RHFGG	gonadoliberin - pi
16	52	30.8	10	1 RHSHG	gonadoliberin - sh
17	52	30.8	89	2 I51423	gonadoliberin prec
18	52	30.8	994	2 S00960	hypothetical prote
19	51	30.2	91	2 JC7393	medaka-type gonado
20	51	30.2	429	2 A54504	circumsporozoite p
21	51	30.2	485	2 A60610	circumsporozoite p
22	51	30.2	507	2 H82378	probable long-chain
23	50	29.6	1484	2 C97196	probable membrane
24	50	29.6	2339	2 A45597	DNA-directed RNA p
25	49.5	29.3	630	2 A36359	VSG expression sit
26	49.5	29.3	630	2 S13724	ESAG 8 protein - t
27	49.5	29.3	999	2 S15961	hypothetical prote
28	49.5	29.3	1265	2 S57968	Ran-binding protei
29	49.5	29.3	3176	1 CGH03A	collagen alpha 3(V

30	49	29.0	532	2 S27373	beta-fructofuranos
31	49	29.0	1146	2 S40311	integrin - fruit f
32	48.5	28.7	630	1 BWUT8Q	regulatory protein
33	48	28.4	10	1 RHAQ1	gonadoliberin I -
34	48	28.4	66	2 T43096	hypothetical prote
35	48	28.4	92	2 I50644	gonadoliberin I pr
36	48	28.4	260	2 T41554	hypothetical prote
37	48	28.4	525	2 T40088	RhodEF domain cont
38	48	28.4	3133	2 S52093	hemocytin - silkwo
39	47.5	28.1	98	2 I50739	gonadotropin-relea
40	47.5	28.1	562	2 JQ1036	63K adducin-like p
41	47.5	28.1	664	2 AB1430	hypothetical prote
42	47.5	28.1	664	2 AH1803	hypothetical prote
43	47.5	28.1	668	1 UYPVME	noncapsid protein
44	47.5	28.1	2244	2 F90563	hypothetical prote
45	47	27.8	217	2 A98196	hypothetical prote

ALIGNMENTS

RESULT 1
A39756
circumsporozoite protein - Plasmodium reichenowi
C:Species: Plasmodium reichenowi
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999
C:Accession: A39756
R:Lal, A.A.; Goldman, I.F.
J. Biol. Chem. 266, 6686-6689, 1991
A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malar
A:Reference number: A39756; MUID:91201303
A:Accession: A39756
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <LAL>
A:Cross-references: GB:M60972; NID:g160228; PIDN:AAA29561.1; PID:g160229
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:312-366/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 48.5%; Score 82; DB 2; Length 388;
Best Local Similarity 78.3%; Pred. No. 0.00084;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIATMEKASVFNVNVGSPSL 24
Db 356 EKKICKMEKCSVFNVNVSSIGL 378

RESULT 2
S05428
circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)
C:Species: Plasmodium falciparum
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000
R:Campbell, J.R.
C:Accession: S05428; A45527; I60657
Nucleic Acids Res. 17, 5854, 1989
A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate
A:Reference number: S05428; MUID:89345189
A:Accession: S05428
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-405 <CAM>
A:Cross-references: EMBL:X15363
R:Caspers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.
Mol. Biochem. Parasitol. 35, 185-190, 1989
A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate
A:Reference number: A45527; MUID:89364998
A:Accession: A45527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <CAS>
A:Cross-references: GB:M22982; GB:J04650; NID:g160168; PIDN:AAA29527.1; PID:g160169
R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.

F:24-92/Product: progonadoliberin #status predicted <PGN>
 F:24-33/Product: gonadoliberin #status predicted <GLN>
 F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
 F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 33.4%; Score 56.5; DB 1; Length 92;
 Best Local Similarity 48.3%; Pred. No. 0.73; Mismatches 12; Indels 1; Gaps 1;
 Matches 14; Conservative 2

QY 5 IARMEKASSVFNNVNSG-PSLHWSYGLRP 32

Db 4 IPKLMAAVLLTVCLGSCSQHWSYGLRP 32

RESULT 7

A44969

C:Species: Plasmodium yoelii nigeriensis

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995

C:Accession: A44969

R:Colomer-Gould, V.; Enea, V.

Mol. Biochem. Parasitol. 43, 51-58, 1990

A:Title: Plasmodium yoelii nigeriensis circumsporozoite gene structure and its implicati

A:Reference number: A44969; MUID:91148645

A:Accession: A44969

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-264 <COL>

A:Cross-references: GB:M32350

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:190-242/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 2; Length 264;

Best Local Similarity 47.4%; Pred. No. 3.9;

Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNNVNS 20

Db 232 DTEICKMDKCSSIFNIVSN 250

RESULT 8

OZZQMB

Circumsporozoite protein precursor - Plasmodium berghei (strain NK65)

N:Alternate names: sporozoite surface antigen

C:Species: Plasmodium berghei

C:Date: 30-Sep-1987 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999

C:Accession: A44948; A25083; S13446

R:Lanar, D.E.

Mol. Biochem. Parasitol. 39, 151-154, 1990

A:Title: Sequence of the circumsporozoite gene of Plasmodium berghei ANKA clone and NK65

A:Reference number: A44948; MUID:90158693

A:Accession: A44948

A:Molecule type: DNA

A:Residues: 1-332 <LAN>

A:Cross-references: GB:M28887

R:Eichinger, D.J.; Arnot, D.E.; Tam, J.P.; Nussenzweig, V.; Enea, V.

Mol. Cell. Biol. 6, 3965-3972, 1986

A:Title: Circumsporozoite protein of Plasmodium berghei: gene cloning and identification

A:Reference number: A25083; MUID:87089740

A:Accession: A25083

A:Molecule type: DNA

A:Residues: 1-26,'I',28-68,'PMLRR',75-126,'P',128-134,'PPPNANDP',135-332 <EITC>

A:Cross-references: GB:M14135; NID:g160245; PIDN:AAA29577.1; PID:g160246

R:Weber, J.L.; Egan, J.E.; Lyon, J.A.; Wirtz, R.A.; Charoenvit, Y.; Maloy, W.L.; Hockmeyer

Exp. Parasitol. 63, 295-300, 1987

A:Title: Plasmodium berghei: cloning of the circumsporozoite protein gene.

A:Reference number: S13446; MUID:87218962

A:Accession: S13446

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 61-122,'A',124-332 <WEB>

A:Cross-references: GB:M25445; NID:g160177; PIDN:AAA29531.1; PID:g160178
 C:Comment: there are three distinct regions in the mature circumsporozoite protein, t
 obic membrane-anchoring sequence.

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

C:Keywords: sporozoite; surface antigen; tandem repeat

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-332/Product: circumsporozoite protein #status predicted <MAT>

F:94-189/Region: 8-residue repeats

F:199-230/Region: 2-residue repeats

F:258-310/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 1; Length 332;

Best Local Similarity 47.4%; Pred. No. 5;

Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNNVNS 20

Db 300 DTEICKMDKCSSIFNIVSN 318

RESULT 9

OZZQMB

Circumsporozoite protein precursor - Plasmodium berghei (strain ANKA clone 2.34L)

N:Alternate names: sporozoite surface antigen

C:Species: Plasmodium berghei

C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999

C:Accession: S07873; S12571

R:Lockyer, M.J.; Davies, C.S.; Suhrbier, A.; Sindén, R.E.

Nucleic Acids Res. 18, 376, 1990

A:Title: Nucleotide sequence of the Plasmodium berghei circumsporozoite protein gene

A:Reference number: S07873; MUID:90221834

A:Accession: S07873

A:Molecule type: DNA

A:Residues: 1-348 <LOC>

A:Cross-references: EMBL:X17606

R:Lockyer, M.J.

submitted to the EMBL Data Library, November 1989

A:Reference number: S12571

A:Accession: S12571

A:Molecule type: DNA

A:Residues: 1-59,'I',61-81,83-348 <LOC2>

A:Cross-references: EMBL:X17606; NID:g9784; PIDN:CAA35608.1; PID:g9785

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

C:Keywords: tandem repeat

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-348/Product: circumsporozoite protein #status predicted <MAT>

F:94-205/Region: 8-residue repeats

F:215-247/Region: 2-residue repeats

F:274-326/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 1; Length 348;

Best Local Similarity 47.4%; Pred. No. 5.3;

Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNNVNS 20

Db 316 DTEICKMDKCSSIFNIVSN 334

RESULT 10

AD2528

Integrase-recombinase protein [imported] - Anabaena sp. (strain PCC 7120) plasmid pCC

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AD2528

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shampo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD2528

A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-360 <KUR>
 A:Cross-references: GB:AP003602; PIDN:BAW77154.1; PID:g17134596; GSPDB:GN00181
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr7511
 A:Genome: plasmid

Query Match 32.5%; Score 55; DB 2; Length 360;
 Best Local Similarity 36.0%; Pred. No. 5.5;
 Matches 9; Conservative 9; Mismatches 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNGPSLHW 26
 I:::|||||: : : : :
 Db 164 EREIAKLEKAAVLDLOHNTNKMHW 188

RESULT 11
 ORZQMY

circumsporozoite protein precursor - Plasmodium yoelii
 N:Alternate names: sporozoite surface antigen
 C:Species: Plasmodium yoelii
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
 C:Accession: A26271
 R:Lal, A.A.; de la Cruz, V.F.; Welsh, J.A.; Charoenvit, Y.; Maloy, W.L.; McCutchan, T.F.
 J. Biol. Chem. 262, 2937-2940, 1987
 A:Title: Structure of the gene encoding the circumsporozoite protein of Plasmodium yoelii
 A:Reference number: A26271; MUID:87137555
 A:Accession: A26271
 A:Molecule type: DNA
 A:Residues: 1-367 <LAW>
 A:Cross-references: GB:J02695; NID:g160222; PIDN:AAA29558.1; PID:g160223
 C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
 C:Comment: There are 15 copies of a 6-residue repeat and 8 copies of a 4-residue repeat.
 C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C:Keywords: sporozoite; surface antigen; tandem repeat
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-367/Product: circumsporozoite protein #status predicted <MAT>
 F:139-228/Region: 6-residue repeats (Q-G-P-G-A-P)
 F:229-260/Region: 4-residue repeats (Q-Q-P-P)
 F:293-345/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 1; Length 367;
 Best Local Similarity 47.4%; Pred. No. 5.6;
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAMEKASSVFNVNS 20
 : : |||: |||: : :
 Db 335 DEICKMDKCSIFNIVSN 353

RESULT 12

gonadolibelin precursor - rhesus macaque (fragment)
 N:Alternate names: luteinizing hormone releasing hormone
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
 C:Accession: I78541
 R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
 Neuroendocrinology 60, 346-359, 1994
 A:Title: Developmental expression of the genes encoding transforming growth factor alpha
 A:Reference number: I58134; MUID:95124501
 A:Accession: I78541
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-67 <RES>
 A:Cross-references: GB:S75918; NID:g912831; PIDN:AA833096.1; PID:g912832

Query Match 32.0%; Score 54; DB 2; Length 67;
 Best Local Similarity 90.0%; Pred. No. 1.2;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32
 I |||||
 Db 5 SQHWSYGLRP 14

RESULT 13

RHMSG
 gonadolibelin precursor - mouse
 N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing hormone
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
 C:Accession: A47578
 R:Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Niko-
 Science 234, 1366-1371, 1986
 A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible
 A:Reference number: A47578; MUID:87069928
 A:Accession: A47578
 A:Molecule type: DNA
 A:Residues: 1-90 <WAS>
 A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
 C:Genetics:
 A:Introns: 45/3; 77/3
 C:Function:
 A:Description: gonadolibelin stimulates pituitary secretion of lutropin and follitropin
 A:Note: gonadolibelin-associated protein may have prolactin release inhibiting activity
 C:Superfamily: gonadolibelin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:22-31/Product: gonadolibelin #status predicted <GLB>
 F:35-90/Product: gonadolibelin-associated protein #status predicted <GAP>
 F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted <GAP>
 F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following-)

Query Match 32.0%; Score 54; DB 1; Length 90;
 Best Local Similarity 90.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32
 I |||||
 Db 21 SQHWSYGLRP 30

RESULT 14

RHHUG
 gonadolibelin precursor [validated] - human
 N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing hormone
 C:Species: Homo sapiens (man)
 C:Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
 C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
 R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
 Nucleic Acids Res 17, 6403-6404, 1989
 A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone
 A:Reference number: S05308; MUID:89366682
 A:Accession: S05308
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-92 <HAY>
 A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
 R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of go a
 A:Reference number: A94090; MUID:86094338
 A:Accession: A26173
 A:Molecule type: mRNA
 A:Residues: 1-92 <ADE>
 A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
 A:Experimental source: hypothalamus
 R:Seeburg, P.H.; Adelman, J.P.
 Nature 311, 666-668, 1984
 A:Title: Characterization of cDNA for precursor of human luteinizing hormone released

A:Reference number: A93342; MUID:85012739
A:Accession: A93342
A:Molecule type: mRNA
A:Residues: 1-15, 'S', 17-92 <SEE>
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
A:Experimental source: placenta
R:Tap, L.; Rousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
A:Reference number: A90108; MUID:83126573
A:Accession: A90108
A:Molecule type: protein
A:Residues: 24-33 <TAN>
A:Experimental source: placental trophoblasts
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterdam
FEBS Lett. 346, 203-206, 1994
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
A:Reference number: S45718; MUID:94283597
A:Contents: annotation; degradation pathway of synthetic hormone
C:Genetics:
A:Gene: GDB:GNRH; LHRH; GRH
A:Cross-references: GDB:I33746; OMIM:227200; OMIM:152760
A:Map position: 8p21-8p11.2
A:Introns: 47/3; 79/3
C:Function:
A:Description: gonadoliberein stimulates pituitary secretion of lutropin and follitropin
A:Note: gonadoliberein-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progonadoliberein #status predicted <PGN>
F:24-33/Product: gonadoliberein #status experimental <MAT>
F:37-92/Product: gonadoliberein-associated protein #status predicted <GAP>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 32.0%; Score 54; DB 1; Length 92;
Best Local Similarity 90.0%; Pred No. 1.7;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 23 SLHWSYGLRP 32
Db 23 SQHWSYGLRP 32
| | | | | | | | | |

RESULT 15
RHPGG
gonadoliberein - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase metho
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A:Note: the synthetic and natural hormones have the same physicochemical and biological
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C:Superfamily: gonadoliberein
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 30.8%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 HWSYGLRP 32
Db 2 HWSYGLRP 9
| | | | | | |

Search completed: October 10, 2002, 16:12:07
Job time : 8.76282 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 5.92308 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-11
Perfect score: 148
Sequence: 1 QYKANSKFIGITELGSLHWSYGLRPX 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119.5	80.7	47	1 US-08-446-692-35	Sequence 35, Appl
2	119.5	80.7	47	2 US-08-488-351A-35	Sequence 35, Appl
3	112.5	76.0	27	1 US-08-446-692-13	Sequence 13, Appl
4	112.5	76.0	27	2 US-08-488-351A-13	Sequence 13, Appl
5	80	54.1	37	1 US-08-446-692-57	Sequence 57, Appl
6	80	54.1	37	2 US-08-488-351A-57	Sequence 57, Appl
7	80	54.1	37	3 US-08-488-351A-63	Sequence 63, Appl
8	80	54.1	37	4 US-08-488-351A-63	Sequence 63, Appl
9	77	52.0	24	5 PCT-US92-07218-25	Sequence 25, Appl
10	77	52.0	24	5 PCT-US95-02121-110	Sequence 110, Appl
11	77	52.0	27	5 PCT-US92-07218-32	Sequence 32, Appl
12	74	50.0	15	2 US-08-319-704-10	Sequence 10, Appl
13	74	50.0	15	2 US-08-661-052-6	Sequence 6, Appl
14	74	50.0	15	2 US-08-460-502-7	Sequence 7, Appl
15	74	50.0	15	4 US-09-046-373-2	Sequence 2, Appl
16	74	50.0	15	4 US-09-188-082-6	Sequence 6, Appl
17	74	50.0	15	5 PCT-US93-11703-69	Sequence 69, Appl
18	74	50.0	16	4 US-09-248-588-55	Sequence 55, Appl
19	74	50.0	17	1 US-08-446-692-4	Sequence 4, Appl
20	74	50.0	17	3 US-08-488-351A-4	Sequence 4, Appl
21	74	50.0	17	3 US-09-100-409A-40	Sequence 40, Appl
22	74	50.0	17	5 PCT-US95-08596-23	Sequence 23, Appl
23	74	50.0	17	5 PCT-US95-13841-7	Sequence 7, Appl
24	74	50.0	31	5 PCT-US93-11703-63	Sequence 63, Appl
25	73	49.3	29	3 PCT-US95-075-257A-13	Sequence 13, Appl
26	73	49.3	29	3 US-09-075-257A-14	Sequence 14, Appl
27	73	49.3	29	4 US-09-534-639-13	Sequence 13, Appl

28	73	49.3	29	4 US-09-534-639-14	Sequence 14, Appl
29	72	48.6	50	4 US-09-171-969-7	Sequence 7, Appl
30	70	47.3	14	1 US-08-186-266-5	Sequence 5, Appl
31	70	47.3	14	1 US-08-305-871A-5	Sequence 5, Appl
32	70	47.3	14	1 US-08-465-167A-18	Sequence 18, Appl
33	70	47.3	14	2 US-08-817-933A-9	Sequence 9, Appl
34	70	47.3	14	5 PCT-US92-07218-15	Sequence 15, Appl
35	70	47.3	14	5 PCT-US92-07218-30	Sequence 30, Appl
36	70	47.3	14	5 PCT-US95-02121-95	Sequence 95, Appl
37	70	47.3	24	5 PCT-US92-07218-31	Sequence 31, Appl
38	70	47.3	27	5 PCT-US92-07218-26	Sequence 26, Appl
39	70	47.3	27	5 PCT-US92-07218-27	Sequence 27, Appl
40	70	47.3	27	5 PCT-US92-07218-28	Sequence 28, Appl
41	70	47.3	27	5 PCT-US95-02121-111	Sequence 111, Appl
42	70	47.3	27	5 PCT-US92-07218-112	Sequence 112, Appl
43	70	47.3	30	5 PCT-US92-07218-29	Sequence 29, Appl
44	70	47.3	32	1 US-08-186-266-9	Sequence 9, Appl
45	69	46.6	15	2 US-08-661-052-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-446-692-35
; Sequence 35, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Y1
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: And synthetic universal immune stimulatoors for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-35

Query Match 80.7%; Score 119.5; DB 1; Length 47;
Best Local Similarity 88.9%; Pred. No. 4.8e-12;
Matches 24; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 QYKANSKFIGITELGSLHWSYGLRP 27
Db 21 QYKANSKFIGITELGGE-HWSYGLRP 46

Fri Oct 11 05:45:01 2002

and synthetic universal immune stimulators for vaccines

```

;
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-13

Query Match 76.0%; Score 112.5; DB 1; Length 27;
Best Local Similarity 85.2%; Pred. No. 3.3e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 1 QYKANSKFIGITELGPSLHWSYGLRP 27
Db 3 QYKANSKFIGITELE---HWSYGLRP 26

RESULT 4
US-08-488-351A-13
; Sequence 13, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-35

Query Match 80.7%; Score 119.5; DB 2; Length 47;
Best Local Similarity 88.9%; Pred. No. 4.8e-12;
Matches 24; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 QYKANSKFIGITELGPSLHWSYGLRP 27
Db 21 QYKANSKFIGITELGGE-HWSYGLRP 46

RESULT 3
US-08-446-692-13
; Sequence 13, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

```

immunity stimulators for vaccines

```

;
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-35

Query Match 80.7%; Score 119.5; DB 2; Length 47;
Best Local Similarity 88.9%; Pred. No. 4.8e-12;
Matches 24; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 QYKANSKFIGITELGPSLHWSYGLRP 27
Db 21 QYKANSKFIGITELGGE-HWSYGLRP 46

RESULT 3
US-08-446-692-13
; Sequence 13, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

```



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-13

Query Match 76.0%; Score 112.5; DB 2; Length 27;
Best Local Similarity 85.2%; Pred. No. 3.3e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 QYIKANSKFIGITELGSLHWSYGLRP 27
Db 3 QYIKANSKFIGITELGSLHWSYGLRP 26

RESULT 5
US-08-446-692-57
; Sequence 57, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-63

Query Match 54.1%; Score 80; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
Db 21 QYIKANSKFIGITELG 36

RESULT 6
US-08-446-692-63
; Sequence 63, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-63

Query Match 54.1%; Score 80; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
Db 5 QYIKANSKFIGITELG 20

RESULT 7
US-08-488-351A-57
; Sequence 57, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-57

```

```

Query Match 54.1%; Score 80; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QYIKANSKFIGITELG 16
Db 21 QYIKANSKFIGITELG 36

```

```

RESULT 8
US-08-488-351A-63
; Sequence 63, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-63

```

```

Query Match 54.1%; Score 80; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QYIKANSKFIGITELG 16
Db 5 QYIKANSKFIGITELG 20

```

```

RESULT 9
PCT-US92-07218-25
; Sequence 25, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitello, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/07218
; FILING DATE: 19920826
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 07/874,491
 FILING DATE: 27-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, William M.
 REGISTRATION NUMBER: 30,223
 REFERENCE/DOCKET NUMBER: 14137-26-3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-543-5043
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 24 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 PCT-US92-07218-25

Query Match 52.0%; Score 77; DB 5; Length 24;
 Best Local Similarity 88.9%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGPS 18
 Db 1 QYIKANSKFIGITEFLPS 18

RESULT 10

PCT-US95-02121-110
 Sequence 110, Application PC/TUS9502121
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
 TITLE OF INVENTION: CTL IMMUNITY
 NUMBER OF SEQUENCES: 153
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/02121
 FILING DATE: 16-FEB-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/197,484
 FILING DATE: 16-FEB-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/935,811
 FILING DATE: 26-AUG-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/874,491
 FILING DATE: 27-APR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/827,682
 FILING DATE: 29-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/749,568
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Patnelee, Steven W.
 REGISTRATION NUMBER: 31,990
 REFERENCE/DOCKET NUMBER: 14137-26-4PC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 467-9600
 TELEFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 110:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 24 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide

PCT-US95-02121-110

Query Match 52.0%; Score 77; DB 5; Length 24;
 Best Local Similarity 88.9%; Pred. No. 1.1e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGPS 18
 Db 1 QYIKANSKFIGITEFLPS 18

RESULT 11

PCT-US92-07218-32
 Sequence 32, Application PC/TUS9207218
 GENERAL INFORMATION:
 APPLICANT: Vitello, Maria A.
 APPLICANT: Chesnut, Robert W.
 TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
 TITLE OF INVENTION: EPITOPES
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend
 STREET: One Market Plaza, Steuart Street Tower
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94105
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/07218
 FILING DATE: 19920826
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/749,568
 FILING DATE: 26-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/827,682
 FILING DATE: 29-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/874,491
 FILING DATE: 27-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, William M.
 REGISTRATION NUMBER: 30,223
 REFERENCE/DOCKET NUMBER: 14137-26-3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-543-5043
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 27 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US92-07218-32

Query Match 52.0%; Score 77; DB 5; Length 27;
 Best Local Similarity 88.9%; Pred. No. 1.3e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGPS 18
 Db 4 QYIKANSKFIGITEFLPS 21

RESULT 12

US-08-319-704-10
 Sequence 10, Application US/08319704

; Patent No. 5814617
; GENERAL INFORMATION:
; APPLICANT: Hoifman, Stephen L.
; APPLICANT: Charoenvit, Yupin
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Doolan, Denise L.
; TITLE OF INVENTION: Protective 17 kDa Malaria Hepatic and
; TITLE OF INVENTION: Erythrocytic Stage Immunogen and Gene
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Naval Medical R & D Command
; STREET: Bldg 1, T-12, 8901 Wisconsin Avenue
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: U.S.A
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/319,704
; FILING DATE: 07-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: A. David Spevack
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: 75,206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 295-6759
; TELEFAX: (301) 295-1022
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-319-704-10

Query Match 50.0%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 13
US-08-661-052-6
; Sequence 6, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-661-052-6

Query Match 50.0%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 14
US-08-460-502-7
; Sequence 7, Application US/08460502
; Patent No. 5843464
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Goirick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-460-502-7

Query Match

50.0%; Score 74; DB 2; Length 15;

```
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

RESULT 15
US-09-046-373-2
; Sequence 2, Application US/09046373
; Patent No. 6235714
; GENERAL INFORMATION:
; APPLICANT: Sudhir Paul
; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Gololobov
; TITLE OF INVENTION: Methods for Identifying Inducers and
; TITLE OF INVENTION: Inhibitors of Catalytic Antibodies, Compositions and Their
; TITLE OF INVENTION: Use
; FILE REFERENCE: UMC 63123
; CURRENT APPLICATION NUMBER: US/09/046,373
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-046-373-2

Query Match 50.0%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
Db 1 QYIKANSKFIGITEL 15

Search completed: October 10, 2002, 16:14:02
Job time : 6.92308 secs
```


PA (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX Claim 8; Page 88; 213pp; English.
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptens containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide of protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC A spacer amino acid sequence (e.g. Gly-Gly) can be provided between the
 CC invasive and Th domains and between the immune stimulator and haptens
 CC components. When the haptens is LHRH, then optionally the invasive domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing immunogenic peptide
 CC as above which can be used as a potent vaccine for treating e.g.
 CC prostatic hyperplasia, androgen-dependent carcinoma, prostatic
 CC carcinoma, testicular carcinoma, endometriosis, benign uterine tumours,
 CC recurrent functional ovarian cysts, (severe) premenstrual syndrome or
 CC oestrogen-dependent breast cancer, or for induction of infertility.
 XX Sequence 47 AA;
 SQ Query Match 80.7%; Score 119.5; DB 15; Length 47;
 Best Local Similarity 88.9%; Pred. No. 3.4e-11; Indels 1; Gaps 1;
 Matches 24; Conservative 0; Mismatches 2;
 QY 1 QYKANSKFIGITELGSLHWSYGLRP 27
 DB 21 QYKANSKFIGITELGGE-HWSYGLRP 46
 RESULT 2
 AAR62701
 ID AAR62701 standard; peptide: 27 AA.
 XX AC AAR62701;
 XX DT 10-SEP-1995 (first entry)
 XX DE LHRH-containing immunogenic peptide.
 XX KW Helper T cell epitope; universal immune stimulator; invasive; haptens;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility;
 KW tetanus toxin.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Domain 1..17
 FT /note= "tetanus toxin helper T cell epitope"
 FT Domain 18..27
 FT /note= "LHRH haptens"
 XX WO9425060-A.
 XX 10-NOV-1994.
 XX 28-APR-1994; 94WO-US04832.
 XX 27-APR-1993; 93US-0057166.
 PR

PR 14-APR-1994; 94US-0229275.
 XX (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX Claims 8, 12; Page 84; 213pp; English.
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptens containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide of protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC A spacer amino acid sequence (e.g. Gly-Gly) can be provided between the
 CC invasive and Th domains and between the immune stimulator and haptens
 CC components. When the haptens is LHRH, then optionally the invasive domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasive-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe) premenstrual
 CC syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX This sequence is particularly preferred.
 XX Sequence 27 AA;
 SQ Query Match 76.0%; Score 112.5; DB 15; Length 27;
 Best Local Similarity 85.2%; Pred. No. 2.2e-10;
 Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
 QY 1 QYKANSKFIGITELGSLHWSYGLRP 27
 DB 3 QYKANSKFIGITELE---HWSYGLRP 26
 RESULT 3
 AAY92639
 ID AAY92639 standard; Protein; 750 AA.
 XX AC AAY92639;
 XX DT 10-AUG-2000 (first entry)
 XX DE Mutant human prostate specific membrane antigen construct, hPSM5.1.
 XX KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX OS Homo sapiens.
 XX Synthetic.
 XX FH Key Location/Qualifiers
 FT Peptide 21..41
 FT /label= P30
 FT /note= "foreign epitope"
 FT 305..319
 FT /label= P2
 FT /note= "foreign epitope"
 XX WO200020027-A2.
 PN


```

PN WO200020027-A2.
XX
PD
XX
XX
XX PF 13-APR-2000.
XX
XX PF 05-OCT-1999; 99WO-DK00525.
XX
XX PR 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX
XX PA (MEBI-) M & E BIOTECH AS.
XX
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX
XX DR WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
XX PT peptide antigens for the treatment of breast and prostate cancer
XX PS
XX PS Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX CC P30). The immunogenic analogues of PSM can be used in the claimed method
XX CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX CC antibody binding regions and cysteine residues involved in disulfide
XX CC bonds are preserved in the immunogenized forms. The method is used for
XX CC inducing immune responses against weakly immunogenic cell-associated
XX CC peptide antigens (PA) such as those associated with cancers
XX CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX CC method comprises effecting simultaneous presentation by antigen producing
XX CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX CC B-cell group derived from the cell-associated PA; and (2) at least 1
XX CC first T helper cell group which is foreign to the animal. Analogues of
XX CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX CC part of all known and predicted CTL and B-cell epitopes of the respective
XX CC PA and including at least one foreign T helper epitope are also claimed.
XX CC The method is used to treat prostate, prostate/breast or breast cancer
XX CC when the PA is human PSM, FGF8b and Her2, respectively.
XX CC Note: This sequence was constructed from the wild type human PSM
XX CC (AAY92619), which appears on pages 184-187 of the specification.
XX
XX SQ Sequence 693 AA;
Query Match 55.4%; Score 82; DB 21; Length 693;
Best Local Similarity 70.4%; Pred. No. 0.00043;
Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

Qy 1 QYKANSKFIGITEL----GPSLHWSY 23
Db 617 QYKANSKFIGITELHVIYAPSSHNY 643
|||||
1 QYKANSKFIGITEL----GPSLHWSY 23
|||||
1 QYKANSKFIGITELHVIYAPSSHNY 700

RESULT 6
AAY92630
ID AAY92630 standard; Protein; 750 AA.
XX
XX AC AAY92630;
XX
XX DT 10-AUG-2000 (first entry)
XX
XX DE Mutant human prostate specific membrane antigen construct, hPSM10.1.
XX
XX KW Prostate specific membrane antigen; immunogenized construct; mutant;
XX KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
XX KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX OS Homo sapiens.
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers

```

```

FT Peptide 21..41
FT /label= P30
FT /note= "foreign epitope"
FT Peptide 674..688
FT /label= P2
FT /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
XX PD 13-APR-2000.
XX
XX XX 05-OCT-1999; 99WO-DK00525.
XX
XX XX 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX
XX XX (MEBI-) M & E BIOTECH AS.
XX
XX XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX
XX XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
XX PT peptide antigens for the treatment of breast and prostate cancer
XX PS
XX PS Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX CC P30). The immunogenic analogues of PSM can be used in the claimed method
XX CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX CC antibody binding regions and cysteine residues involved in disulfide
XX CC bonds are preserved in the immunogenized forms. The method is used for
XX CC inducing immune responses against weakly immunogenic cell-associated
XX CC peptide antigens (PA) such as those associated with cancers
XX CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX CC method comprises effecting simultaneous presentation by antigen producing
XX CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX CC B-cell group derived from the cell-associated PA; and (2) at least 1
XX CC first T helper cell group which is foreign to the animal. Analogues of
XX CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX CC part of all known and predicted CTL and B-cell epitopes of the respective
XX CC PA and including at least one foreign T helper epitope are also claimed.
XX CC The method is used to treat prostate, prostate/breast or breast cancer
XX CC when the PA is human PSM, FGF8b and Her2, respectively.
XX CC Note: This sequence was constructed from the wild type human PSM
XX CC (AAY92619), which appears on pages 184-187 of the specification.
XX
XX SQ Sequence 750 AA;
Query Match 55.4%; Score 82; DB 21; Length 750;
Best Local Similarity 70.4%; Pred. No. 0.00047;
Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

Qy 1 QYKANSKFIGITEL----GPSLHWSY 23
Db 674 QYKANSKFIGITELHVIYAPSSHNY 700
|||||
1 QYKANSKFIGITEL----GPSLHWSY 23
|||||
1 QYKANSKFIGITELHVIYAPSSHNY 700

RESULT 7
AAY92641
ID AAY92641 standard; Protein; 750 AA.
XX
XX AC AAY92641;
XX
XX DT 10-AUG-2000 (first entry)
XX
XX DE Mutant human prostate specific membrane antigen construct, hPSM10.0.
XX KW Prostate specific membrane antigen; immunogenized construct; mutant;

```

KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 674..688
 FT /label= P2
 FT /note= "foreign epitope"
 FT
 FT W0200020027-A2.
 XX
 XX 13-APR-2000.
 XX
 XX 05-OCT-1999; 99WO-DK00525.
 XX
 XX 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX
 PA (MEBI-) M & E BIOTECH AS.
 XX
 XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX
 XX WPI; 2000-349917/30.
 DR
 XX
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 XX Example 1; Page -: 220pp; English.
 XX
 CC AAY92627-49 are mutant immunogenized human prostate specific membrane
 CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
 CC P30). The immunogenic analogues of PSM can be used in the claimed method,
 CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
 CC antibody binding regions and cysteine residues involved in disulfide
 CC bonds are preserved in the immunogenized forms. The method is used for
 CC inducing immune responses against weakly immunogenic cell-associated
 CC peptide antigens (PA) such as those associated with cancers
 CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
 CC method comprises effecting simultaneous presentation by antigen producing
 CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
 CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the respective
 CC PA and including at least one foreign T helper epitope are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer.
 CC when the PA is human PSM, FGF8b and Her2, respectively.
 CC Note: This sequence was constructed from the wild type human PSM
 CC (AAY92619), which appears on pages 184-187 of the specification.
 XX
 SQ Sequence 750 AA;
 Query Match 55.4%; Score 82; DB 21; Length 750;
 Best Local Similarity 70.4%; Pred. No. 0.00047;
 Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
 QY 1 QYIKANSKFIGITEL----GPSLHWSY 23
 |||||
 Db 674 QYIKANSKFIGITELHVIYAPSSHNY 700
 RESULT 8
 AAY92646
 ID AAY92646 standard; Protein; 750 AA.
 XX
 AC AAY92646;
 XX
 DT 10-AUG-2000 (first entry)

XX
 DE Mutant human prostate specific membrane antigen construct, hPSM10.3.
 XX
 KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 674..688
 FT /label= P2
 FT /note= "foreign epitope"
 FT
 FT Peptide 210..230
 FT /label= P30
 FT /note= "foreign epitope"
 XX
 XX W0200020027-A2.
 XX
 XX 13-APR-2000.
 XX
 XX 05-OCT-1999; 99WO-DK00525.
 XX
 XX 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX
 PA (MEBI-) M & E BIOTECH AS.
 XX
 XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX
 XX WPI; 2000-349917/30.
 DR
 XX
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 XX Example 1; Page -: 220pp; English.
 XX
 CC AAY92627-49 are mutant immunogenized human prostate specific membrane
 CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
 CC P30). The immunogenic analogues of PSM can be used in the claimed method,
 CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
 CC antibody binding regions and cysteine residues involved in disulfide
 CC bonds are preserved in the immunogenized forms. The method is used for
 CC inducing immune responses against weakly immunogenic cell-associated
 CC peptide antigens (PA) such as those associated with cancers
 CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
 CC method comprises effecting simultaneous presentation by antigen producing
 CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
 CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the respective
 CC PA and including at least one foreign T helper epitope are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer.
 CC when the PA is human PSM, FGF8b and Her2, respectively.
 CC Note: This sequence was constructed from the wild type human PSM
 CC (AAY92619), which appears on pages 184-187 of the specification.
 XX
 SQ Sequence 750 AA;
 Query Match 55.4%; Score 82; DB 21; Length 750;
 Best Local Similarity 70.4%; Pred. No. 0.00047;
 Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
 QY 1 QYIKANSKFIGITEL----GPSLHWSY 23
 |||||
 Db 674 QYIKANSKFIGITELHVIYAPSSHNY 700

Db 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31

DE tetanus toxoid T cell epitope and der p11 B cell epitope peptide.
XX
KW T cell epitope; B cell epitope; allergy; allergen; antigenic;
KW antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW atopic dermatitis; acute urticaria; chronic urticaria

KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
 KW anaphylactic reaction; drug hypersensitivity; allergic reaction.
 XX
 OS Dermatophagoides pteronyssinus.
 OS Clostridium tetani.
 OS Synthetic.
 XX
 PN W0200006694-A2.
 XX
 XX 10-FEB-2000.
 XX
 XX 20-JUL-1999; 99WO-BE000092.
 XX
 PR 30-JUL-1998; 98EP-0870167.
 XX
 PA (UNIO) UCB SA.
 XX
 XX Saint-Remy J, Jacquemin M;
 XX WPI: 2000-422470/36.
 DR
 XX
 XX New compound for prevention and treatment of allergies comprises at
 PT least one allergen antigenic determinant recognized by a B cell and at
 PT least one antigenic determinant which does not trigger T cell
 PT activation.
 XX
 PS Claim 8; Page 35; 50pp; English.
 XX
 XX The present invention describes a compound (I) for the prevention and/or
 CC treatment of allergy. The compound comprises at least one allergen
 CC antigenic determinant (i) recognised by a B cell or an antibody secreted
 CC by a B cell of a non-atopic individual and at least one antigenic
 CC determinant (ii) different from the allergen that triggers T cell
 CC activation. (I) has antiallergic, antiasthmatic, antiinflammatory,
 CC dermatological and immunosuppressive activities, and can be used in a
 CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
 CC treat and/or prevent allergies or a disease of allergic origin,
 CC especially hypersensitivities. These include rhinitis, sinusitis,
 CC bronchial asthma, atopic dermatitis, some forms of acute and chronic
 CC urticaria, gastro-intestinal syndromes associated with the ingestion of
 CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions
 CC associated with drug hypersensitivities and/or a mixture of these. The
 CC use of (I) in the treatment of allergic conditions avoids the need for
 CC drug treatment, which often causes undesirable side-effects. Also, prior
 CC art drug therapies alleviate symptoms, but do not influence their
 CC causes, however (I) actually combats the cause of an allergic reaction.
 CC The present sequence represents a specifically claimed compound peptide
 CC sequence from the present invention.
 XX
 SQ Sequence 32 AA;
 Query Match 54.1%; Score 80; DB 21; Length 32;
 Best Local Similarity 100.0%; Pred. No. 2.8e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITELG 16
 Db | | | | | | | | | | | | | | | |
 1 QYIKANSKFIGITELG 16
 RESULT 12
 AAR65389
 ID AAR65389 standard; peptide; 37 AA.
 XX
 AC AAR65389;
 XX
 DT 21-SEP-1995 (first entry)
 XX
 DE Universal immunostimulator having GG spacers.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; haptens;
 KW tetanus toxin.
 XX

OS Synthetic.
 XX Key Location/Qualifiers
 FH Domain 3..19
 FT /note= "tetanus toxin helper T cell epitope"
 FT Domain 22..37
 FT /note= "invasin domain"
 XX
 PN W09425060-A.
 XX
 XX 10-NOV-1994.
 PD
 XX
 XX 28-APR-1994; 94WO-US04832.
 PF
 XX
 PR 27-APR-1993; 93US-0057166.
 PR 14-APR-1994; 94US-0229275.
 XX
 XX (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 XX
 XX Ladd AE, Wang CY, Zamb T;
 PI WPI: 1994-357910/44.
 DR
 XX
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 PT
 XX
 PS Disclosure; Page 95; 213pp; English.
 XX
 XX Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptens containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasin and Th domains and between the immune stimulator and haptens
 CC components. When the haptens is LHRH, then optionally the invasin domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence is an example of a -GG-Th-GG-invasin immune
 CC stimulator to which a haptens can be bonded.
 XX
 SQ Sequence 37 AA;
 Query Match 54.1%; Score 80; DB 15; Length 37;
 Best Local Similarity 100.0%; Pred. No. 3.3e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITELG 16
 Db | | | | | | | | | | | | | | | |
 5 QYIKANSKFIGITELG 20
 RESULT 13
 AAR65383
 ID AAR65383 standard; peptide; 37 AA.
 XX
 AC AAR65383;
 XX
 DT 21-SEP-1995 (first entry)
 XX
 DE Universal immunostimulator having GG spacers.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; haptens;
 KW tetanus toxin.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Domain 1..16

DE Tetanus toxoid T cell epitopes and Der pII B cell epitopes peptide.
 XX
 KW T cell epitope; B cell epitope; allergy; allergen; antigenic;
 KW anti-allergic; antiasthmatic; antiinflammatory; dermatological;
 KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
 KW atopic dermatitis; acute urticaria; chronic urticaria;
 KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
 KW anaphylactic reaction; drug hypersensitivity; allergic reaction.
 XX
 OS Dermatophagoides pteronyssinus.
 OS Clostridium tetani.
 OS Synthetic.
 XX
 PN WO200006694-A2.
 XX
 PD 10-FEB-2000.
 XX
 PF 20-JUL-1999; 99WO-BE00092.
 XX
 PR 30-JUL-1998; 98EP-0870167.
 XX
 PA (UNIO) UCB SA.
 XX
 XX Saint-Remy J, Jacquemin M;
 PI
 DR WPI; 2000-422470/36.
 XX
 PT New compound for prevention and treatment of allergies comprises at
 PT least one allergenic determinant recognized by a B cell and at
 PT least one antigenic determinant which does not trigger T cell
 PT activation -
 XX
 XX Claim 8; Page 35; 50pp; English.
 XX
 CC The present invention describes a compound (I) for the prevention and/or
 CC treatment of allergy. The compound comprises at least one allergen
 CC antigenic determinant (i) recognised by a B cell or an antibody secreted
 CC by a B cell of a non-atopic individual and at least one antigenic
 CC determinant (ii) different from the allergen that triggers T cell
 CC activation. (i) has anti-allergic, antiasthmatic, antiinflammatory,
 CC dermatological and immunosuppressive activities, and can be used in a
 CC vaccine. (i) may be used in a pharmaceutical or cosmetic medicament to
 CC treat and/or prevent allergies or a disease of allergic origin,
 CC especially hypersensitivities. These include rhinitis, sinusitis,
 CC bronchial asthma, atopic dermatitis, some forms of acute and chronic
 CC urticaria, gastro-intestinal syndromes associated with the ingestion of
 CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions
 CC associated with drug hypersensitivities and/or a mixture of these. The
 CC use of (i) in the treatment of allergic conditions avoids the need for
 CC drug treatment, which often causes undesirable side-effects. Also, prior
 CC art drug therapies alleviate symptoms, but do not influence their
 CC causes, however (i) actually combats the cause of an allergic reaction.
 CC The present sequence represents a specifically claimed compound peptide
 CC sequence from the present invention.
 XX
 SQ Sequence 137 AA;
 Query Match 54.1%; Score 80; DB 21; Length 137;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKEFIGITELG 16
 DB 2 QYIKANSKEFIGITELG 17
 |||||
 Search completed: October 10, 2002, 16:05:09
 Job time : 15.7179 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 10.8889 Seconds

(without alignments)

444.844 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKFITGLPGLHWSYGLRXP 28

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTEMBL19.*
 2: sp_archaea.*
 3: sp_bacteria.*
 4: sp_fungi.*
 5: sp_human.*
 6: sp_invertebrate.*
 7: sp_mammal.*
 8: sp_mhc.*
 9: sp_organelle.*
 10: sp_phage.*
 11: sp_plant.*
 12: sp_rodent.*
 13: sp_virus.*
 14: sp_vertebrate.*
 15: sp_unclassified.*
 16: sp_rvirus.*
 17: sp_bacteriap.*
 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	50.0	1310	2 Q93N27	Q93n27 clostridium
2	52.5	35.5	342	16 Q98FZ0	Q98fz0 rhizobium l
3	52	35.1	90	13 Q90Y63	Q90y63 rana catesb
4	52	35.1	91	13 Q9PRH0	Q9prh0 anguilla ja
5	52	35.1	382	5 Q95VY3	Q95vy3 lumbricus t
6	51	34.5	384	5 Q77072	Q77072 eisenia foe
7	51	34.5	388	17 Q9YD14	Q9ydl4 aeropyrum p
8	50.5	34.1	60	7 Q31585	Q31585 salmo salar
9	50.5	34.1	71	7 Q9XRJ9	Q9xri9 salvelinus
10	50.5	34.1	85	7 Q95IS2	Q95is2 salmo salar
11	50.5	34.1	85	7 Q95HY1	Q95hy1 salmo salar
12	50.5	34.1	86	7 Q95HX4	Q95hx4 salmo salar
13	50.5	34.1	244	7 Q31590	Q31590 salmo salar
14	50	33.8	220	3 O14264	O14264 schizosacch
15	50	33.8	480	16 Q9I482	Q9i482 pseudomonas
16	49.5	33.4	67	7 Q31578	Q31578 salmo salar

17	49	33.1	324	12 Q89914	Q89914 sonchus yel
18	49	33.1	484	16 Q9HZ30	Q9hz30 pseudomonas
19	48	32.4	2091	3 P78616	P78616 pmericella
20	47.5	32.1	84	13 Q9DEK4	Q9dek4 coregonus s
21	47.5	32.1	85	7 Q95IS3	Q95is3 salmo salar
22	47.5	32.1	85	7 Q95IR2	Q95ir2 salmo salar
23	47.5	32.1	149	7 Q31495	Q31495 oncorhynch
24	47.5	32.1	205	17 Q981D4	Q981d4 sulfolobus
25	47.5	32.1	216	7 Q9GJH0	Q9gjh0 salmo trutt
26	47.5	32.1	216	7 Q9GJG9	Q9gig9 salmo trutt
27	47	31.8	134	10 Q9LSN9	Q9lsn9 arabidopsis
28	47	31.8	728	16 Q9HY06	Q9hyq6 pseudomonas
29	46.5	31.4	67	7 Q31577	Q31577 salmo salar
30	46.5	31.4	67	7 Q31581	Q31581 salmo salar
31	46.5	31.4	67	7 Q31582	Q31582 salmo salar
32	46.5	31.4	85	7 Q95IS5	Q95is5 salmo salar
33	46.5	31.4	85	7 Q95HX9	Q95hx9 salmo salar
34	46.5	31.4	85	7 Q95HX8	Q95hx8 salmo salar
35	46.5	31.4	85	7 Q95HX6	Q95hx6 salmo salar
36	46.5	31.4	85	7 Q95HX5	Q95hx5 salmo salar
37	46.5	31.4	85	7 Q95HX3	Q95hx3 salmo salar
38	46.5	31.4	193	16 Q99T58	Q99t58 staphylococ
39	46.5	31.4	245	7 Q31591	Q31591 salmo salar
40	46.5	31.4	342	7 O19442	O19442 mus musculu
41	46.5	31.4	388	16 Q9KM21	Q9km21 vibrio chol
42	46	31.1	56	8 O20728	O20728 wasmannia a
43	46	31.1	87	13 Q9YI26	Q9yi26 sparus aura
44	46	31.1	142	2 Q9R749	Q9r749 coxiella bu
45	46	31.1	171	2 Q9S619	Q9s619 coxiella bu

ALIGNMENTS

RESULT 1

Q93N27
 ID Q93N27 PRELIMINARY; PRT; 1310 AA.

AC Q93N27; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TETANUS TOXIN (FRAGMENT).
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shumin Z., Dianliang L.;
 RT "Cloning and sequence analysis of tetanus toxin gene.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF389424; AAK72964.2; -.
 FT NON_TER 1
 FT NON_TER 1310
 SQ SEQUENCE 1310 AA; 150316 MW; 9EADD914418E450 CRC64;

Query Match 50.0%; Score 74; DB 2; Length 1310;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFITGL 15
 Db 831 QYIKANSKFITGL 845

RESULT 2

Q98FZ0
 ID Q98FZ0 PRELIMINARY; PRT; 342 AA.

AC Q98FZ0; 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE MLL3560 PROTEIN.

OS	Anguilla japonica (Japanese eel).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Anguilloidei;
OC	Anguillidae; Anguilla.
OX	NCBI_TaxID=7937;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BRAIN;
RA	Okubo K., Suetake H., Aida K.;
RT	"A splicing variant for the prepro-mammalian gonadotropin-releasing
RT	hormone (prepro-mGnRH) mRNA is present in the brain and various
RT	peripheral tissues of the Japanese eel.";
RL	Zool. Sci. 16:645-651(1999).
RL	[2]
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC	SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR	EMBL; AB026989; BAA82608.1; -.
DR	EMBL; AB026991; BAA83597.1; -.
DR	InterPro; IPR002012; GnRH.
DR	InterPro; IPR004079; Gonadoliberin1.
DR	Pfam; PF00446; GnRH; 1.
DR	PRINTS; PR01541; GONADOLIBRN1.
DR	PROSITE; PS00473; GnRH; 1.
KW	Annotation: Hormone; Signal.
FT	SIGNAL 1 22
FT	CHAIN 23 32
FT	CHAIN 33 91
FT	SEQUENCE 91 AA; 9893 MW; BA15C9DC08434A7B CRC64;
QY	Query Match 35.1%; Score 52; DB 13; Length 91;
QY	Best Local Similarity 100.0%; Pred. No. 1.7;
QY	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps
QY	20 HWSYGLRP 27
QY	24 HWSYGLRP 31
QY	11111111
RESULT 5	
ID5VY3	ID Q95VY3 PRELIMINARY; PRT; 382 AA.
AC	Q95VY3;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	COELOMIC CYTOLYTIC FACTOR PRECURSOR.
DE	Lumbricus terrestris (Common earthworm).
OS	Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC	Lumbricina; Lumbricidae; Lumbricus.
OX	NCBI_TaxID=6398;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Beschin A., De Baetselier P., Bilej M.;
RT	"Distinct carbohydrate recognition domains of an earthworm defense
RT	molecule recognize Gram negative and Gram positive bacteria.";
RT	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL	EMBL; AF395805; AAL09587.1; -.
RL	Signal.
KW	SIGNAL 1 17
KW	SIGNAL 17
FT	POTENTIAL.
FT	SEQUENCE 382 AA; 43931 MW; 5256CF171EB7D3FB CRC64;
QY	Query Match 35.1%; Score 52; DB 5; Length 382;
QY	Best Local Similarity 36.8%; Pred. No. 8.5;
QY	Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps
QY	6 NSKFIGITELGPSLHWSYG 24
QY	11111111

```
Db 196 SGEPGIGKMGSTMHGPG 214
RESULT 6
O77072
AC 077072 PRELIMINARY; PRT; 384 AA.
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE COELOMIC CYTOLYTIC FACTOR 1.
GN CCF1.
OS Euisenia foetida (Common brandling worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Euisenia.
OX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98406152; PubMed=9733802;
RA Beschin A., Bilej M., Hanssens F., Raymakers J., Van Dyck E.,
RA Revets H., Brys L., Gomez J., De Baetselier P., Timmermans M.;
RT "Identification and cloning of a glucan- and Lipopolysaccharide-
RT binding protein from Euisenia foetida earthworm involved in the
RT activation of prophenoloxidase cascade.";
RL J. Biol. Chem. 273:24948-24954 (1998).
DR EMBL: AF030028; AAC35887.1;
DR InterPro: IPR000757; Glyco_hydro_16.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR SEQ 384 AA; 44322 MW; C90B5C94003BAD6D CRC64;
Query Match 34.5%; Score 51; DB 5; Length 384;
Best Local Similarity 41.2%; Pred. No. 12;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 8 KFIGITELPSLHWSYG 24
:||||:|:|:|
Db 198 EFLGIQKMGSTMHGPG 214
RESULT 7
O9YD14
ID O9YD14 PRELIMINARY; PRT; 388 AA.
AC O9YD14;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 388AA LONG HYPOTHETICAL FMU PROTEIN.
GN APE1098.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT Crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101 (1999).
DR EMBL: AP000060; BAA80083.1;
DR InterPro: IPR001678; Noll_Nop2_Sun.
DR InterPro: IPR002478; PUA.
DR Pfam: PF01189; Noll_Nop2_Sun; 1.
DR Pfam: PF01472; PUA; 1.
DR SMART: SM00359; PUA; 1.
KW Complete proteome.
SQ SEQUENCE 388 AA; 42535 MW; A9E10CEAAEF0B0AC CRC64;
```

```
Query Match 34.5%; Score 51; DB 5; Length 388;
Best Local Similarity 41.4%; Pred. No. 12;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 14 ELGPSLHWSYGLRP 27
||| | |||
Db 345 ELGRLTWSWGLRP 358
||| | |||
RESULT 8
Q31585
ID Q31585 PRELIMINARY; PRT; 60 AA.
AC Q31585;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE (DB02) MHC CLASS II BETA 1 (FRAGMENT).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Grimholt U., Olsaker I., de Vries Lindstrom C., Lie O.;
RT "A study of polymorphism in the MHC class II beta 1 and MHC class I
RT alpha 2 domain exons of Atlantic salmon (Salmo salar).";
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: L24953; AAA49597.1;
DR InterPro: IPR000353; MHC_II_beta.
DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6776 MW; 5D4F3449060940E2 CRC64;
Query Match 34.1%; Score 50.5; DB 7; Length 60;
Best Local Similarity 57.9%; Pred. No. 1.9;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;
QY 1 QYIKANS---KFIGITELG 16
:||||:|:|:|
Db 16 EYIRFNTVGRFVGYTELG 34
:||||:|:|:|
RESULT 9
O9XRJ9
ID O9XRJ9 PRELIMINARY; PRT; 71 AA.
AC O9XRJ9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MHC CLASS II BETA 1 (FRAGMENT).
GN SANA.
OS Salvelinus namaycush (lake trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8040;
RN [1]
RP SEQUENCE FROM N.A.
RA Dorschner M.O., Duris T., Phillips R.B.;
RT "Diversity of a Lake Trout Mhc class II Gene.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF130026; AAD20889.1;
DR InterPro: IPR000353; MHC_II_beta.
DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 7176 MW; 5D4F3449060940E2 CRC64;
```

SO SEQUENCE 71 AA; 8002 MW; E3095286582A9F2D CRC64;

Query Match 34.1%; Score 50.5; DB 7; Length 71;
Best Local Similarity 57.9%; Pred. No. 2.3;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 QYIKANS---KFIGITELG 16
:||: || ||: || ||||
Db 14 EYIRFNSTVGKFGVGYTELG 32

RESULT 10

Q95IS2 ID Q95IS2 PRELIMINARY; PRT; 85 AA.

AC Q95IS2; DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC CLASS II BETA CHAIN (FRAGMENT).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Landry C., Bernatchez L.;
RT "Comparative analysis of population structure across environments and
RT geographic scales at Major Histocompatibility Complex and
RT Microsatellite in Atlantic salmon (Salmo salar).";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373699; AAK61882.1; -
KW MHC.
FT NON_TER 1 1
FT NON_TER 85 85
SQ SEQUENCE 85 AA; 9743 MW; 3214E01AD1B66AC5 CRC64;

Query Match 34.1%; Score 50.5; DB 7; Length 85;
Best Local Similarity 57.9%; Pred. No. 2.8;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 QYIKANS---KFIGITELG 16
:||: || ||: || ||||
Db 33 EYIRFNSTVGKFGVGYTELG 51

RESULT 11

Q95HY1 ID Q95HY1 PRELIMINARY; PRT; 85 AA.

AC Q95HY1; DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC CLASS II B ANTIGEN (FRAGMENT).
GN DB1.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Landry C., Bernatchez L.;
RT "Allelic polymorphism in MHC class II B in four populations of
RT Atlantic salmon (Salmo salar).";
RL Immunogenetics 53:329-336(2001).
DR EMBL; AF104370; AAL04002.1; -
KW MHC.
FT NON_TER 1 1
FT NON_TER 85 85
SQ SEQUENCE 85 AA; 9723 MW; 27C9F7931F1F01C4 CRC64;

Query Match 34.1%; Score 50.5; DB 7; Length 85;
Best Local Similarity 57.9%; Pred. No. 2.8;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 QYIKANS---KFIGITELG 16
:||: || ||: || ||||
Db 33 EYIRFNSTVGKFGVGYTELG 51

RESULT 12

Q95HX4 ID Q95HX4 PRELIMINARY; PRT; 86 AA.

AC Q95HX4; DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC CLASS II B ANTIGEN (FRAGMENT).
GN DB1.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Landry C., Bernatchez L.;
RT "Allelic polymorphism in MHC class II B in four populations of
RT Atlantic salmon (Salmo salar).";
RL Immunogenetics 53:329-336(2001).
DR EMBL; AF104377; AAL04009.1; -
KW MHC.
FT NON_TER 1 1
FT NON_TER 86 86
SQ SEQUENCE 86 AA; 9912 MW; E5097729F681F149 CRC64;

Query Match 34.1%; Score 50.5; DB 7; Length 86;
Best Local Similarity 57.9%; Pred. No. 2.8;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 QYIKANS---KFIGITELG 16
:||: || ||: || ||||
Db 33 EYIRFNSTVGKFGVGYTELG 51

RESULT 13

Q31590 ID Q31590 PRELIMINARY; PRT; 244 AA.

AC Q31590; DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC CLASS II.
GN MHC-SASA CLASS II B.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Hardwick I., Grimholt U., Fosse V.M., Lie Y., Endresen C.;
RT "Cloning and sequence analysis of cDNAs encoding the MHC class II a-
RT chain in Atlantic salmon, Salmo salar.";
RL Immunogenetics 37:437-441(1993).
DR EMBL; X70166; CAA49725.1; -
DR HSSP; P01888; 1BMG.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00047; ig; 1.

```
DR Pfam: PF00969; MHC_II_beta: 1.
DR ProDom: PD000328; MHC_II_beta; 1.
DR SMART: SM00407; IGcl; 1.
KW Glycoprotein; MHC II; Transmembrane.
SQ SEQUENCE 244 AA; 27449 MW; 496CB9EA9D73765C CRC64;

Query Match          34.1%; Score 50.5; DB 7; Length 244;
Best Local Similarity 57.9%; Pred. No. 8.9;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 OYIKANS---KFIGITELG 16
   :||: || ||: || ||||
Db 51 EYIRNSTVGKFGVYTELG 69

RESULT 14
O14264
ID O14264 PRELIMINARY; PRT; 220 AA.
AC O14264;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 25.8 KDA PROTEIN C7D4.09C IN CHROMOSOME I.
GN SPAC7D4.09C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBSJ databases.
DR EMBL; Z99532; CAB16726.2; -.
DR InterPro; IPR001104; S5A_redtse_C.
KW Hypothetical protein.
SQ SEQUENCE 220 AA; 25760 MW; 8314536BD00595C8 CRC64;

Query Match          33.8%; Score 50; DB 3; Length 220;
Best Local Similarity 29.4%; Pred. No. 9.5;
Matches 10; Conservative 7; Mismatches 9; Indels 8; Gaps 1;

QY 2 YIKANSKFIGITELGPSLHW-----SYGLRP 27
   | : : | : | : ||| |
Db 18 YFTSTFLVFSILKNAPSLSLWLMKYGHDNFGDKP 51

RESULT 15
Q9I482
ID Q9I482 PRELIMINARY; PRT; 480 AA.
AC Q9I482;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PROBABLE MFS TRANSPORTER.
GN PA1262.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
```

```
RL Nature 406:959-964(2000).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AE004556; AAG04651.1; -.
KW Complete proteome; Transmembrane.
SQ SEQUENCE 480 AA; 50407 MW; 4E05690D837B86C5 CRC64;

Query Match          33.8%; Score 50; DB 16; Length 480;
Best Local Similarity 88.9%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 LGPSLHWSY 23
   |||||
Db 224 LGPSLHWSW 232

Search completed: October 10, 2002, 16:09:58
Job time : 12.8889 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:19:47 ; Search time 16 seconds
(without alignments)
60.056 Million cell updates/sec

Title: US-09-848-834A-1
Sequence: 1 XHSYGLRPX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 1039

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	52	96.3	10	1 RHPGC	gonadoliberein - pig
2	52	96.3	10	1 RUSHG	gonadoliberein - sh
3	48	88.9	10	1 RHAQ1	gonadoliberein I -
4	39	72.2	10	2 A21114	gonadoliberein I - ch
5	34	63.0	10	1 RHAQ2	gonadoliberein II -
6	34	63.0	10	1 A61126	gonadoliberein - sp
7	34	63.0	10	2 A46030	gonadoliberein I -
8	34	63.0	10	2 B46030	gonadoliberein II -
9	31	57.4	10	2 A49187	gonadotropin-relea
10	25	46.3	10	2 P00177	neuromedin C - lau
11	25	46.3	10	2 A60647	fatty-acid synthas
12	19	35.2	7	2 A60139	neuropeptide calla
13	19	35.2	8	2 D47393	tachykinin I - mig
14	19	35.2	10	1 ECLQ1M	gonadoliberein III - m
15	19	35.2	10	1 ECLQ3M	Ig heavy chain CRD
16	18	33.3	9	2 P02999	gonadoliberein - se
17	18	33.3	10	1 RHLMGS	hypotrehalosemic h
18	18	33.3	10	2 B33995	Ig mu chain J regi
19	17	31.5	8	2 PH1618	Ig H chain V-D-J r
20	17	31.5	9	2 S07205	litorin 2-Glu - Au
21	17	31.5	9	2 S07204	litorin I - Austr
22	17	31.5	9	2 DS8503	translation elonga
23	17	31.5	9	2 P02398	Ig heavy chain CRD
24	17	31.5	9	2 PH1591	Ig H chain V-D-J r
25	17	31.5	9	2 G41946	T-cell receptor ga
26	17	31.5	10	2 A49581	sialokinin I - yel
27	17	31.5	10	2 B49581	sialokinin II - yel
28	17	31.5	10	2 A24867	scyllokinin I - s
29	17	31.5	10	2 A24867	scyllokinin I - s

ALIGNMENTS

RESULT 1

RHPGC gonadoliberein - pig

C:Species: Sus scrofa domestica (domestic pig)
C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411

R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971

A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 822-827, 1971

A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me

A:Reference number: A90176; MUID:72065376

A:Contents: annotation; synthesis

A:Molecule type: protein

A:Residues: 1-10 <BAB>

R:Baba, Y.; Arimura, A.; Nair, R.M.G.; Schally, A.V.

Biochem. Biophys. Res. Commun. 45, 483-487, 1971

A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.

A:Reference number: A90175; MUID:72117544

A:Contents: annotation

A:Note: Trp-3 appears to be essential for biological activity

C:Superfamily: gonadoliberein

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyroglutamic acid (Gln) #status experimental

F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 96.3%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0017;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9

DB 2 HWSYGLRP 9

RESULT 2

RHSHG

gonadoliberein - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier,

Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteliniz-

A:Reference number: A93780; MUID:72094314

A:Molecule type: protein

Fri Oct 11 06:06:03 2002

A;Residues: 1-10 <BUR>
 A;Note: the natural and synthetic hormones have the same biological activity
 C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormones

C;Superfamily: gonadoliberein
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 96.3%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0017; Indels 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0

QY 2 HWSYGLRP 9
 |||||
 Db 2 HWSYGLRP 9

RESULT 3

RHAQ1
 gonadoliberein I - American alligator
 N;Alternate names: gonadotropin-releasing hormone I
 C;Species: Alligator mississippiensis (American alligator)
 C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
 C;Accession: A60066

R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson, R.; Regal, Pept. 33, 105-116, 1991
 A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of Alligator mississippiensis

A;Reference number: A60066; MUID:91352338

A;Accession: A60066
 A;Molecule type: protein
 A;Residues: 1-10 <LOV>
 C;Superfamily: gonadoliberein
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 88.9%; Score 48; DB 1; Length 10;
 Best Local Similarity 87.5%; Pred. No. 0.0087; Indels 0; Gaps 0;
 Matches 7; Conservative 1; Mismatches 0

QY 2 HWSYGLRP 9
 |||||
 Db 2 HWSYGLRP 9

RESULT 4

A21114
 gonadoliberein - chum salmon
 C;Species: Oncorhynchus keta (chum salmon)
 C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
 C;Accession: A21114

R;Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.; Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983

A;Title: Characterization of a teleost gonadotropin-releasing hormone.

A;Reference number: A21114; MUID:83195140

A;Accession: A21114

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <SHE>

Query Match 72.2%; Score 39; DB 2; Length 10;
 Best Local Similarity 75.0%; Pred. No. 0.36; Indels 2; Gaps 0;
 Matches 6; Conservative 0; Mismatches 2

QY 2 HWSYGLRP 9
 |||||
 Db 2 HWSYGLRP 9

RESULT 5:

RHAQ2
 gonadoliberein II - American alligator
 N;Alternate names: gonadotropin-releasing hormone II

C;Species: Alligator mississippiensis (American alligator)

C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
 C;Accession: B60066

R;Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan, Regul. Pept. 33, 105-116, 1991

A;Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of Alligator mississippiensis

A;Reference number: A60066; MUID:91352338

A;Accession: B60066

A;Molecule type: protein

A;Residues: 1-10 <LOV>
 C;Superfamily: gonadoliberein
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 63.0%; Score 34; DB 1; Length 10;
 Best Local Similarity 62.5%; Pred. No. 2.9; Indels 2; Gaps 0;
 Matches 5; Conservative 1; Mismatches 2

QY 2 HWSYGLRP 9
 |||||
 Db 2 HWSYGLRP 9

RESULT 6

A61126
 gonadoliberein - spotted ratfish

N;Alternate names: gonadotropin-releasing hormone
 C;Species: Hydroloagus collii (spotted ratfish)

C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
 C;Accession: A61126

R;Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.; Gen. Comp. Endocrinol. 82, 152-161, 1991

A;Title: Primary structure of gonadotropin-releasing hormone from the brain of a holo-

A;Reference number: A61126; MUID:91340067

A;Accession: A61126

A;Molecule type: protein

A;Residues: 1-10 <LOV>

C;Superfamily: gonadoliberein

C;Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gly) #status experimental

F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 63.0%; Score 34; DB 1; Length 10;
 Best Local Similarity 62.5%; Pred. No. 2.9; Indels 2; Gaps 0;
 Matches 5; Conservative 1; Mismatches 2

QY 2 HWSYGLRP 9
 |||||
 Db 2 HWSYGLRP 9

RESULT 7

A46030
 gonadoliberein I - spiny dogfish

N;Alternate names: gonadotropin-releasing hormone

C;Species: Squalus acanthias (spiny dogfish)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998

R;Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992

A;Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain

A;Reference number: A46030; MUID:92335300

A;Accession: A46030

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <LOV>

C;Keywords: hormone; pyroglutamic acid (Gln) #status experimental

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 63.0%; Score 34; DB 2; Length 10;
 Best Local Similarity 62.5%; Pred. No. 2.9;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 Db 2 HWSHGWL 9

RESULT 8
 B46030
 gonadoliberin II - spiny dogfish
 N:Alternate names: gonadotropin-releasing hormone
 C:Species: Squalus acanthias (spiny dogfish)
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
 C:Accession: B46030
 R:Lovejoy, D.A.; Fischer, W.H.; Ngamvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
 Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
 A:Title: Distinct sequence of gonadotropin-releasing hormone (GNRH) in dogfish brain pro
 A:Reference number: A46030; MUID:92335300
 A:Accession: B46030
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <LOW>
 C:Keywords: hormone; pyroglutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 63.0%; Score 34; DB 2; Length 10;
 Best Local Similarity 62.5%; Pred. No. 2.9;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 Db 2 HWSHGWL 9

RESULT 9
 A49187
 gonadotropin-releasing hormone III - sea lamprey
 C:Species: Petromyzon marinus (sea lamprey)
 C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
 C:Accession: A49187
 R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
 Endocrinology 132, 1125-1131, 1993
 A:Title: Primary structure and biological activity of a third gonadotropin-releasing hor
 A:Reference number: A49187; MUID:93178316
 A:Accession: A49187
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <SOW>
 A:Experimental source: brain
 A>Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 57.4%; Score 31; DB 2; Length 10;
 Best Local Similarity 50.0%; Pred. No. 9.9;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
 Db 2 HWSHDWKP 9

RESULT 10
 P00177
 neuromedin C - laughing frog
 C:Species: Rana ridibunda (laughing frog)
 C>Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 11-Jan-2000
 C:Accession: P00177
 R:Conlon, J.M.; O'Harte, F.; Vaudry, H.
 Biochem. Biophys. Res. Commun. 178, 526-530, 1991
 A:Title: Primary structures of the bombesin-like neuropeptides in frog brain show that b
 A:Reference number: P00177; MUID:91315477
 A:Accession: P00177
 A:Molecule type: protein
 A:Residues: 1-10 <CON>

A:Experimental source: brain
 C:Superfamily: gastrin-releasing peptide
 C:Keywords: amidated carboxyl end
 F:10/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 46.3%; Score 25; DB 2; Length 10;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 6
 Db 3 HWSYGLRP 7

RESULT 11
 A60647
 neuromedin C - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 07-May-1999
 C:Accession: A60647
 R:Lemaire, S.; Trifaro, J.M.; Chouinard, L.; Cecyre, D.; Dessureault, J.; Mercier, P.
 Peptides 10, 355-360, 1989
 A:Title: Structural identification, subcellular localization and secretion of bovine
 A:Reference number: A60647; MUID:89331342
 A:Accession: A60647
 A:Molecule type: protein
 A:Residues: 1-10 <LEM>
 A>Note: this neuropeptide was purified from secretory granules of cells in the adrena
 C:Superfamily: gastrin-releasing peptide
 C:Keywords: adrenal gland; neuropeptide

Query Match 46.3%; Score 25; DB 2; Length 10;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRP 6
 Db 3 HWSYGLRP 7

RESULT 12
 A60139
 fatty-acid synthase (EC 2.3.1.85) - rabbit (fragment)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-May-2000
 C:Accession: A60139
 R:Hardie, D.G.; Dewart, K.B.; Aitken, A.; McCarthy, A.D.
 Biochim. Biophys. Acta 828, 380-382, 1985
 A:Title: Amino acid sequence around the reactive serine residue of the thioesterase d
 A:Reference number: A60139; MUID:85175165
 A:Accession: A60139
 A:Molecule type: protein
 A:Residues: 1-7 <HAR>
 C:Superfamily: rat fatty-acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I h
 ydrolase homology; short-chain alcohol dehydrogenase homology; [acyl-carrier-protein]
 C:Keywords: acyltransferase; carrier protein; coenzyme A; homodimer; multifunctional
 F:5/Active site: Ser (of oleoyl-[acyl-carrier-protein] hydrolase) #status experimental

Query Match 35.2%; Score 19; DB 2; Length 7;
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSYGLRP 6
 Db 4 WSYGLRP 7

RESULT 13
 D47393
 neuropeptide callatostatin 4 - bluebottle fly (Calliphora vomitoria)
 C:Species: Calliphora vomitoria
 C>Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change
 C:Accession: D47393

Fri Oct 11 06:06:03 2002

Search completed: October 10, 2002, 16:39:44
Job time : 17 secs

R:Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.
Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993
A:Title: Callatostations: neuropeptides from the blowfly Calliphora vomitoria with sequen
A:Reference number: A47393; MUID:93211980
A:Accession: D47393
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <DUV>
A:Experimental source: thoracic ganglia
A>Note: sequence extracted from NCBI backbone (NCBIP:128479)

Query Match 35.2%; Score 19; DB 2; Length 8;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 WSYGL 7
DB 4 YSFGL 8

RESULT 14

ECLOIM
tachykinin I - migratory locust
C:Species: Locusta migratoria (migratory locust)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 08-Dec-1995
C:Accession: S08265
R:Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; de Loof, A.
FEBS Lett. 261, 397-401, 1990
A:Title: Locustatachykinin I and II, two novel insect neuropeptides with homology to pep
A:Reference number: S08265; MUID:90184489
A:Accession: S08265
A:Molecule type: protein
A:Residues: 1-10 <SCH>
C:Superfamily: tachykinin
C:Keywords: amidated carboxyl end; neuropeptide; tachykinin
F:10/Modified site: amidated carboxyl end (Arg) #status experimental

Query Match 35.2%; Score 19; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 YGLR 8
DB 7 YGVR 10

RESULT 15

ECLOIM
tachykinin III - migratory locust
N:Alternate names: locustatachykinin III
C:Species: Locusta migratoria (migratory locust)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 08-Dec-1995
C:Accession: A60073
R:Schoofs, L.; Holman, G.M.; Hayes, T.K.; Kochansky, J.P.; Nachman, R.J.; De Loof, A.
Regul. Pept. 31, 199-212, 1990
A:Title: Locustatachykinin III and IV: two additional insect neuropeptides with homology
A:Reference number: A60073; MUID:91219696
A:Accession: A60073
A:Molecule type: protein
A:Residues: 1-10 <SCH>
C:Superfamily: tachykinin
C:Keywords: amidated carboxyl end; neuropeptide; tachykinin
F:10/Modified site: amidated carboxyl end (Arg) #status experimental

Query Match 35.2%; Score 19; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 YGLR 8
DB 7 YGVR 10

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:12:26 ; Search time 13 Seconds
(without alignments)
29.784 Million cell updates/sec

Title: US-09-848-834A-1
Perfect score: 54
Sequence: 1 XHSYGLRXP 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 349

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	48	88.9	10	1 GON1_ALLMI	P37041 alligator m
2	41	75.9	10	1 GON1_CLUPA	P81749 clupea pall
3	39	72.2	10	1 GON3_ONCKE	P20367 oncorhynch
4	34	63.0	10	1 GON2_CHICK	P37043 gallus gall
5	34	63.0	10	1 GONL_SQUAC	P27429 squalus aca
6	31	57.4	10	1 GON3_PETMA	P30948 petromyzon
7	26	48.1	10	1 GON1_CHEPR	P80677 chelyosoma
8	25	46.3	10	1 GRP_RANRI	P32260 rana ridibu
9	24	44.4	8	1 ALL1_CYPDO	P82152 cydia pomon
10	23	42.6	8	1 ALI16_CARMA	P81819 carcinus ma
11	23	42.6	10	1 GON2_CHEPR	P80678 chelyosoma
12	19	35.2	5	1 ALI14_CARMA	P81817 carcinus ma
13	19	35.2	8	1 ALI15_CARMA	P81818 carcinus ma
14	19	35.2	8	1 ALI17_CARMA	P81820 carcinus ma
15	19	35.2	8	1 ALI18_CARMA	P81821 carcinus ma
16	19	35.2	8	1 ALL3_CYPDO	P82154 cydia pomon
17	19	35.2	8	1 ALL4_CALVO	P81840 calliphora
18	19	35.2	8	1 ALL4_CYPDO	P82155 cydia pomon
19	19	35.2	9	1 TKC1_CALVO	P41517 calliphora
20	19	35.2	9	1 TKL1_LOEMI	P46223 locusta mig
21	19	35.2	10	1 ALI19_CARMA	P81822 carcinus ma
22	19	35.2	10	1 TKL2_LOEMI	P46224 locusta mig
23	19	35.2	10	1 TKL3_LOEMI	P30249 locusta mig
24	18	33.3	10	1 GON1_PETMA	P43378 petromyzon
25	18	33.3	10	1 HTE_TABAT	P44596 tabanus atr
26	17	31.5	9	1 LITO_LITAU	P08945 litoria aur
27	17	31.5	10	1 TKN1_SCYCA	P08608 scyllorhinu
28	17	31.5	10	1 TKNB_RANCA	P22689 rana catesb
29	17	31.5	10	1 TKS1_AEDAE	P42634 aedes aegypt
30	17	31.5	10	1 TKS2_AEDAE	P42635 aedes aegypt
31	16	29.6	7	1 ALL2_CARMA	P81805 carcinus ma
32	16	29.6	7	1 ALL3_CARMA	P81806 carcinus ma
33	16	29.6	7	1 ALL4_CARMA	P81807 carcinus ma

34	16	29.6	7	1	ALI5_CARMA	P81808 carcinus ma
35	16	29.6	8	1	ALI2_CARMA	P81815 carcinus ma
36	16	29.6	8	1	ALI6_CYPDO	P82157 cydia pomon
37	16	29.6	8	1	ALL7_CARMA	P81809 carcinus ma
38	16	29.6	8	1	ALL8_CARMA	P81811 carcinus ma
39	16	29.6	8	1	ALI9_CARMA	P81812 carcinus ma
40	16	29.6	9	1	ALI10_CARMA	P81813 carcinus ma
41	16	29.6	9	1	ALI11_CARMA	P81814 carcinus ma
42	16	29.6	9	1	LITR_PHYRO	P08946 phyllomedus
43	16	29.6	10	1	TRP6_LEUMA	P81739 leucophaea
44	16	29.6	10	1	TRP7_LEUMA	P82158 cydia pomon
45	15	27.8	7	1	ALL7_CYPDO	

ALIGNMENTS

RESULT 1

GON1_ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I)
DE (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 88.9%; Score 48; DB 1; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.003;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYGLRP 9
|||||
Db 2 HWSYGLQP 9

RESULT 2

GON1_CLUPA STANDARD; PRT; 10 AA.
AC P81749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH)
DE (Luliberin I).
GN GNRH1.
OS Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;

OC Clupeinae; Clupea.
 OX NCBI_TaxID=30724;
 RN [1]
 RP SEQUENCE, AND FUNCTION.
 RC TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
 Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring."
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
 DR InterPro: IPR002012; GnRH.
 DR Pfam: PF00446; GnRH; 1.
 DR PROSITE: PS00473; GnRH; 1.
 KW Hormone: Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;
 Query Match 75.9%; Score 41; DB 1; Length 10;
 Best Local Similarity 75.0%; Pred. No. 0.059;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HWSYGLRP 9
 |||:| |
 DB 2 HWSHGLSP 9
 |||:| |
 RESULT 3
 GON3_ONCKE STANDARD; PRT; 10 AA.
 AC P20367; P81751;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadoliberein III (Gonadotropin-releasing hormone III) (GnRH-III) (LH-
 RH III) (Luliberin III).
 GN GNRH3.
 OS Oncorhynchus keta (Chum salmon), and
 OS Clupea pallasii (Pacific herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8018, 30724;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=O.keta;
 RX MEDLINE=83195140; PubMed=6341999;
 RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
 RT "Characterization of a teleost gonadotropin-releasing hormone."
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
 RN [2]
 RP SEQUENCE, AND FUNCTION.
 RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
 Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring."
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 FOLLICLE-STIMULATING HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
 DR PIR; A21114; A21114.
 DR InterPro: IPR002012; GnRH.
 DR Pfam: PF00446; GnRH; 1.
 DR PROSITE: PS00473; GnRH; 1.
 KW Hormone: Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.

FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;
 Query Match 72.2%; Score 39; DB 1; Length 10;
 Best Local Similarity 75.0%; Pred. No. 0.14;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 HWSYGLRP 9
 |||:| |
 DB 2 HWSYGLWP 9
 |||:| |
 RESULT 4
 GON2_CHICK STANDARD; PRT; 10 AA.
 AC P37043; P20408; P81750;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadoliberein II (Gonadotropin-releasing hormone II) (GnRH-II)
 (LH-RH II) (Luliberin II).
 OS Gallus gallus (Chicken).
 OS Alligator mississippiensis (American alligator).
 OS Squalus acanthias (Spiny dogfish).
 OS Hydrolagus collii (Spotted ratfish) (Pacific ratfish), and
 OS Clupea pallasii (Pacific herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Chicken; TISSUE=Hypothalamus;
 RX MEDLINE=84222059; PubMed=6427779;
 RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
 Matsuo H.;
 RT "Identification of the second gonadotropin-releasing hormone in
 chicken hypothalamus: evidence that gonadotropin secretion is
 RT probably controlled by two distinct gonadotropin-releasing hormones
 in avian species."
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=A.mississippiensis; TISSUE=Brain;
 RX MEDLINE=91335238; PubMed=1882082;
 RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
 Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure of two forms of gonadotropin-releasing hormone
 from brains of the American alligator (Alligator mississippiensis)."
 RL Regul. Pept. 33:105-116(1991).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=S.acanthias; TISSUE=Brain;
 RX MEDLINE=92333300; PubMed=1631133;
 RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
 Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
 RT "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
 RT dogfish brain provides insight into GnRH evolution."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=H.collii; TISSUE=Brain;
 RX MEDLINE=91340067; PubMed=1678723;
 RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
 Lee T.;
 RT "Primary structure of gonadotropin-releasing hormone from the brain
 RT of a holoccephalan (ratfish: Hydrolagus collii)."
 RL Gen. Comp. Endocrinol. 82:152-161(1991).
 RN [5]
 RP SEQUENCE, AND FUNCTION.
 RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

RA Chang J.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring.";
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR PIR: B60066; RHA02.
 DR PIR: A61126; A61126.
 DR PIR: B46030; B46030.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1254 MW; 284B2E437871F5A3 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 10;
 Best Local Similarity 62.5%; Pred. No. 1.2;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 HWSYGLRP 9
 ||||:|
 Db 2 HWSHGWP 9

RESULT 5
 GONL_SQUAC STANDARD; PRT; 10 AA.
 ID GONL_SQUAC
 AC P27429;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Gonadoliberin (Gonadotropin-releasing hormone) (GNRH) (LH-RH)
 DE (Luliberin).
 OS Squalus acanthias (Spiny dogfish).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 CC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
 OX NCBI_TaxID=7797;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=92333300; PubMed=1631133;
 RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
 RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
 RT "Distinct sequence of gonadotropin-releasing hormone (GNRH) in
 RT dogfish brain provides insight into GNRH evolution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 10;
 Best Local Similarity 62.5%; Pred. No. 1.2;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 HWSYGLRP 9
 ||||:|
 Db 2 HWSHGWP 9

RESULT 6
 GON3_PETMA STANDARD; PRT; 10 AA.
 ID GON3_PETMA
 AC P30948;

DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III)
 DE (Luliberin III).
 OS Petromyzon marinus (Sea lamprey).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 CC Petromyzontiformes; Petromyzontidae; Petromyzon.
 OX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=93178316; PubMed=8440174;
 RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
 RT "Primary structure and biological activity of a third gonadotropin-
 RT releasing hormone from lamprey brain.";
 RL Endocrinology 132:1125-1131(1993).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR Pfam: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1277 MW; 284B36237A1F5A3 CRC64;

Query Match 57.4%; Score 31; DB 1; Length 10;
 Best Local Similarity 50.0%; Pred. No. 4.1;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 2 HWSYGLRP 9
 ||||:|
 Db 2 HWSHDWKP 9

RESULT 7
 GONL_CHEPR STANDARD; PRT; 10 AA.
 ID GONL_CHEPR
 AC P80677;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I)
 DE (Luliberin I).
 OS Chelyosoma productum.
 CC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
 CC Corellidae; Chelyosoma.
 OX NCBI_TaxID=71177;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96413669; PubMed=8816823;
 RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
 RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
 RT "Two new forms of gonadotropin-releasing hormone in a protochordate
 RT and the evolutionary implications.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
 CC THE GONADUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
 CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
 CC -!- MASS SPECTROMETRY: MW=1246.56; METHOD=MALDI.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR InterPro: IPR002012; GNRH.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation.
 FT MOD_RES 1 1
 FT MOD_RES 10 10
 SQ SEQUENCE 10 AA; 1264 MW; 284B3639DB5AB5A3 CRC64;

```
Query Match 48.1%; Score 26; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
   ||| :|
DB 2 HWSDFPKP 9

RESULT 8
GRP_RANRI
ID GRP_RANRI STANDARD; PRT; 10 AA.
AC P23260;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neuromedin C.
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8406;
RN [1]
RP SEQUENCE
RC TISSUE=Brain;
RX MEDLINE=91315477; PubMed=1859413;
RA Conlon J.M., O'Harte F., Vaudry H.;
RT "Primary structures of the bombesin-like neuropeptides in frog brain
RL Biochem Biophys Res Commun. 178:526-530(1991).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
CC FAMILY.
DR PIR; PQ0177; PQ0177.
DR InterPro; IPR000874; Bombesin.
DR Pfam; PF02044; Bombesin; 1.
DR PROSITE; PS00257; BOMBESIN; 1.
KW Bombesin family; Amidation.
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1094 MW; F81FBAE862CDC371 CRC64;

Query Match 46.3%; Score 25; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 53;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYG 6
   ||: |
DB 3 HWAAG 7

RESULT 9
ALLI_CYPDO
ID ALLI_CYPDO STANDARD; PRT; 8 AA.
AC P82152;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiaastatin 1.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.

FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 934 MW; C82879C45B51F775 CRC64;

Query Match 44.4%; Score 24; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGL 7
   ||::||
DB 3 HYNFGL 8

RESULT 10
ALI6_CARMA
ID ALI6_CARMA STANDARD; PRT; 8 AA.
AC P81819;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 16.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Fortunoidea; Fortunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 813 MW; 7C286B45AB476878 CRC64;

Query Match 42.6%; Score 23; DB 1; Length 8;
Best Local Similarity 80.0%; Pred. No. 1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSYGL 7
   :|||
DB 4 YSYGL 8

RESULT 11
GON2_CHEPR
ID GON2_CHEPR STANDARD; PRT; 10 AA.
AC P80678;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GnRH-II)
DE (Luliberin II).
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
RA Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
and the evolutionary implications.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
```

CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
CC THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -1- MASS SPECTROMETRY: MW=1117.52; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; GNRH. 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 6 6 INTERCHAIN.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
SQ MOD_RES 10 AA; 1135 MW; 284B38D1EEB735A3 CRC64;

Query Match 42.6%; Score 23; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWS 4
Db 1 1 1
2 HWS 4

RESULT 12

AL14_CARMA
ID AL14_CARMA STANDARD; PRT; 5 AA.
AC P81817;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 14.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]

SEQUENCE.

RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
CC Neuropeptide; Amidation; Multigene family.
KW MOD_RES 5 5 AMIDATION (POTENTIAL).
FT MOD_RES 5 5
SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 35.2%; Score 19; DB 1; Length 5;
Best Local Similarity 60.0%; Pred. No. 1e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSYGL 7
Db 1 1 1
YSFGL 5

RESULT 13

AL15_CARMA
ID AL15_CARMA STANDARD; PRT; 8 AA.
AC P81818;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 15.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;

RN SEQUENCE.
RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
CC Neuropeptide; Amidation; Multigene family.
KW MOD_RES 8 8 AMIDATION.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 811 MW; 922879D5AB47687D CRC64;

Query Match 35.2%; Score 19; DB 1; Length 8;
Best Local Similarity 60.0%; Pred. No. 1e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSYGL 7
Db 4 1 1
YSFGL 8

RESULT 14

AL17_CARMA
ID AL17_CARMA STANDARD; PRT; 8 AA.
AC P81820;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 17.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]

SEQUENCE.

RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
CC Neuropeptide; Amidation; Multigene family.
KW MOD_RES 8 8 AMIDATION (POTENTIAL).
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 35.2%; Score 19; DB 1; Length 8;
Best Local Similarity 60.0%; Pred. No. 1e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSYGL 7
Db 4 1 1
YSFGL 8

RESULT 15

AL18_CARMA
ID AL18_CARMA STANDARD; PRT; 8 AA.
AC P81821;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 18.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.

OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 8 8 AMIDATION (POTENTIAL).
 SQ SEQUENCE 8 AA; 919 MW; C82879D5AB569AB5 CRC64;

Query Match 35.2%; Score 19; DB 1; Length 8;
 Best Local Similarity 60.0%; Pred. No. ie+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSYGL 7
 Db 4 YSFGL 8

Search completed: October 10, 2002, 16:38:18
 Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:14:17 ; Search time 28 Seconds
(without alignments)
61.784 Million cell updates/sec

Title: US-09-848-834A-1
Perfect score: 54
Sequence: 1 XHWSYGLRPX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.19.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rviro.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	20	37.0	10	8 Q958J8	Q958J8 rana muscosa
2	19	35.2	8	6 O02831	O02831 oryctolagus
3	18	33.3	8	8 Q94VC1	Q94VC1 varanus rud
4	17	31.5	8	13 P79940	P79940 xenopus lae
5	17	31.5	10	2 Q47561	Q47561 escherichia
6	17	31.5	10	10 Q99213	Q99213 aegilops sq
7	17	31.5	10	11 Q9ESU5	Q9ESU5 mus musculus
8	16	29.6	10	11 Q9QVE6	Q9QVE6 mus sp. pro
9	15	27.8	8	4 Q15888	Q15888 homo sapien
10	15	27.8	8	8 Q94VB5	Q94VB5 varanus sal
11	15	27.8	8	8 Q94VB2	Q94VB2 varanus sal
12	15	27.8	8	8 Q94VA7	Q94VA7 varanus sal
13	15	27.8	8	8 Q94PX7	Q94PX7 felis silve
14	15	27.8	8	8 Q94PX6	Q94PX6 felis libyc
15	15	27.8	8	8 Q94PX5	Q94PX5 felis silve
16	15	27.8	8	11 P82598	P82598 rattus norv

17	15	27.8	8	12	064971	064971 alfalfa mos
18	15	27.8	9	4	Q9BYF9	Q9BYF9 homo sapien
19	15	27.8	9	8	Q94VC6	Q94VC6 varanus pil
20	15	27.8	9	11	062530	062530 mus spretus
21	15	27.8	9	12	065711	065711 berne virus
22	15	27.8	9	13	Q9PRJ4	Q9PRJ4 leptososteus
23	15	27.8	10	8	Q9XMB4	Q9XMB4 aegilops ta
24	15	27.8	10	8	Q94VD5	Q94VD5 varanus oli
25	15	27.8	10	12	069347	069347 herpes simp
26	14.5	26.9	8	2	085406	085406 coxiella bu
27	14	25.9	10	15	086324	086324 rous sarcom
28	14	25.9	10	15	086325	086325 rous sarcom
29	14	25.9	10	15	086326	086326 rous sarcom
30	13.5	25.0	8	13	Q98TU5	Q98TU5 xenopus lae
31	13	24.1	7	10	049223	049223 glycine max
32	13	24.1	8	2	009258	009258 synechococc
33	13	24.1	8	2	052062	052062 bacillus me
34	13	24.1	9	5	Q9VW82	Q9VW82 drosophila
35	13	24.1	9	7	Q31415	Q31415 gallus gall
36	13	24.1	9	8	Q94XE6	Q94XE6 tectocoris
37	13	24.1	9	8	Q94NB2	Q94NB2 microcebus
38	13	24.1	9	8	Q94NB1	Q94NB1 microcebus
39	13	24.1	9	8	Q94NB0	Q94NB0 microcebus
40	13	24.1	9	8	Q94NA9	Q94NA9 daubentonia
41	13	24.1	10	2	Q48469	Q48469 kiebsteilia
42	13	24.1	10	4	Q96QA7	Q96QA7 homo sapien
43	13	24.1	10	6	Q9TS43	Q9TS43 sus scrofa
44	13	24.1	10	8	Q94VD2	Q94VD2 varanus pan
45	13	24.1	10	11	Q9QVE5	Q9QVE5 mus sp. pro

ALIGNMENTS

RESULT 1

Q958J8 PRELIMINARY; PRT; 10 AA.
AC Q958J8
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN COI.
OS Rana muscosa.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=160500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21184280; PubMed=11286498;
RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
RA Jennings M., Larsson A.;
RT "Molecular Phylogenetics of Western North American Frogs of the Rana
boylii Species Group.";
RL Mol. Phylogenet. Evol. 19:131-143(2001).
DR EMBL; AF314026; AAK56898.1;
KW Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1335 MW; COD380C9D371F1A9 CRC64;

Query Match 37.0%; Score 20; DB 8; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 HWSY 5
|| :
Db 5 HWFF 8

RESULT 2

O02831 PRELIMINARY; PRT; 8 AA.
ID O02831

```

AC 002831;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PRO ALPHA 1 TYPE III COLLAGEN PROTEIN (FRAGMENT).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96377339; PubMed=8783186;
RA Metzaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA Vuorio E.;
RT "Evidence for insufficient chondrocytic differentiation during repair
RT of full-thickness defects of articular cartilage.";
RL Matrix Biol. 15:39-47(1996).
DR EMBL; S83371; AAD14433.1; -.
KW Collagen.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 1028 MW; B859C7272EA77371 CRC64;

Query Match 35.2%; Score 19; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HW 3
DB 1 HW 2

RESULT 3
ID Q94VC1 PRELIMINARY; PRT; 8 AA.
AC Q94VC1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN COI.
OS Varanus rudicollis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scieroglossa; Anguilliformia; Varanidae; Varanus.
OX NCBI_TaxID=169851;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407521; AAL10116.1; -.
KW Mitochondrion.
FT NON_TER 8
SQ SEQUENCE 8 AA; 1053 MW; FE2729D5A36411A6 CRC64;

Query Match 33.3%; Score 18; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSY 5
DB 4 WSF 6

RESULT 4
ID P79940 PRELIMINARY; PRT; 8 AA.
AC P79940;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE XMEIS1-4 PROTEIN (FRAGMENT).
OS Xenopus laevis (African clawed frog).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97202105; PubMed=9049632;
RA Steelman S., Moskow J.J., Muzynski K., North C., Druck T.,
RA Montgomery J.C., Huebner K., Daar I.O., Buchberg A.M.;
RT "Identification of a conserved family of Meis1-related homeobox
RT genes.";
RL Genome Res. 7:142-156(1997).
DR EMBL; U68389; AAB19199.1; -.
DR TRANSFAC; T03410; -.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 1187 MW; 278B51F37B11F40B CRC64;

Query Match 31.5%; Score 17; DB 13; Length 8;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WSY 5
DB 5 WHY 7

RESULT 5
ID Q47561 PRELIMINARY; PRT; 10 AA.
AC Q47561;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE HYPOTHETICAL 1.1 KDA PROTEIN (FRAGMENT).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K-12;
RX MEDLINE=94162733; PubMed=7764507;
RA Yamada M., Yanai S., Talkuder A.;
RT "Analysis of products of the Escherichia coli genomic genes and
RT regulation of their expressions: an applicable procedure for genomic
RT analysis of other microorganisms.";
RL Biosci. Biotechnol. Biochem. 58:117-120(1994).
DR EMBL; D21143; BRA04679.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1109 MW; 2D1B58B1E87DD733 CRC64;

Query Match 31.5%; Score 17; DB 2; Length 10;
Best Local Similarity 37.5%; Pred. No. 5.8e+03;
Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
DB 2 NWLGHSP 9

RESULT 6
ID Q99213 PRELIMINARY; PRT; 10 AA.
AC Q99213;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE ALBUMIN (FRAGMENT).
OS Aegilops squarrosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;

```

```

OC Triziceae; Aegilops.
RX NCBI_TaxID=37682;
RN [1]
RP SEQUENCE.
RA Shewry P.R., Lafandra D., Salcedo G., Aragoncillo C.,
RL Garcia-Olmedo F., Lew E.J.-L., Dietler M.D., Kasarda D.D.;
RA FEBS Lett. 175:359-363(1984).
KW Seed storage protein.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1105 MW; 3A1AB5AEA365A367 CRC64;

Query Match 31.5%; Score 17; DB 10; Length 10;
Best Local Similarity 66.7%; Pred. No. 5.8e+03;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 WSY 5
Db 4 WSW 6

RESULT 7
Q9ESU5
ID Q9ESU5 PRELIMINARY; PRT; 10 AA.
AC Q9ESU5
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE FAS DEATH RECEPTOR (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HE;
RX MEDLINE=20127858; PubMed=10660538;
RA Munsch D., Watanabe-Fukunaga R., Bourdon J.C., Nagata S., May E.,
RA Yonish-Rouach E., Relsdorf P.;
RT "Human and mouse Fas (APO-1/CD95) death receptor genes each contain a
RT p53-responsive element that is activated by p53 mutants unable to
RT induce apoptosis.";
RL J. Biol. Chem. 275:3867-3872(2000).
DR EMBL; AF282865; AAG02410.1; -.
KW Receptor.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1242 MW; 22145E32CDC37043 CRC64;

Query Match 31.5%; Score 17; DB 11; Length 10;
Best Local Similarity 28.6%; Pred. No. 5.8e+03;
Matches 2; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 WSYGLRP 9
Db 3 WIWAVLP 9

RESULT 8
Q9QVE6
ID Q9QVE6 PRELIMINARY; PRT; 10 AA.
AC Q9QVE6
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE PROTAMINE MP2 INTERMEDIATE PROTEIN PMP2/26.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=92174934; PubMed=1541289;
RA Chauviere M., Martinage A., Debarle M., Sautiere P., Chevallier P.;
RT "Molecular characterization of six intermediate proteins in the

```

```

RT "processing of mouse protamine p2 precursor.";
RL Eur. J. Biochem. 204:759-765(1992).
SQ SEQUENCE 10 AA; 1028 MW; 2B099C75B72866D8 CRC64;

Query Match 29.6%; Score 16; DB 11; Length 10;
Best Local Similarity 75.0%; Pred. No. 8.8e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GLRP 9
Db 5 GLSP 8

RESULT 9
Q15888
ID Q15888 PRELIMINARY; PRT; 8 AA.
AC Q15888
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE (CLONE XP15H8A) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinnault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32069; AAA73878.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 1068 MW; 0315A37EAB5B0763 CRC64;

Query Match 27.8%; Score 15; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WS 4
Db 6 WS 7

RESULT 10
Q94VB5
ID Q94VB5 PRELIMINARY; PRT; 8 AA.
AC Q94VB5
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
OS COI.
OC Varanus salvator cumingi.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169830;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407523; AAL10122.1; -.
KW Mitochondrion.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 992 MW; EFC775A5A36411A6 CRC64;

Query Match 27.8%; Score 15; DB 8; Length 8;

```

Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WS 4

Db 4 WS 5

RESULT 11

Q94VB2 PRELIMINARY; PRT; 8 AA.
AC Q94VB2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN COI.
OS Varanus salvator togianus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169832;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RL "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407524; AAL10125.1; -.
KW Mitochondrion.
FT NON_TER 8
SQ SEQUENCE 8 AA; 992 MW; EFC775A5A36411A6 CRC64;

Query Match 27.8%; Score 15; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WS 4

Db 4 WS 5

RESULT 12

Q94VA7 PRELIMINARY; PRT; 8 AA.
AC Q94VA7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN COI.
OS Varanus salvator salvator.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169831;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RL "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407526; AAL10130.1; -.
KW Mitochondrion.
FT NON_TER 8
SQ SEQUENCE 8 AA; 992 MW; EFC775A5A36411A6 CRC64;

Query Match 27.8%; Score 15; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WS 4

Db 4 WS 5

RESULT 13

Q94PX7 PRELIMINARY; PRT; 8 AA.
AC Q94PX7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT II (FRAGMENT).
GN COII.
OS Felis silvestris catus (Cat).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1, 2, 7, 12, 16, 17, AND 110;
RA Randi E., Pierpaoli M., Beaumont M., Ragni B., Sforzi A.;
RT "Genetic identification of wild and domestic cats (Felis silvestris), and their hybrids using Bayesian clustering methods.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ409128; CAC41027.1; -.
DR EMBL; AJ409129; CAC41030.1; -.
DR EMBL; AJ409130; CAC41033.1; -.
DR EMBL; AJ409131; CAC41036.1; -.
DR EMBL; AJ409132; CAC41039.1; -.
DR EMBL; AJ409133; CAC41042.1; -.
DR EMBL; AJ409134; CAC41045.1; -.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 8 AA; 951 MW; 262685BDC5A3733B CRC64;

Query Match 27.8%; Score 15; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WS 4

Db 3 WS 4

RESULT 14

Q94PX6 PRELIMINARY; PRT; 8 AA.
AC Q94PX6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT II (FRAGMENT).
GN COII.
OS Felis libyca.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=61377;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=40, 1, 2, AND 7;
RA Randi E., Pierpaoli M., Beaumont M., Ragni B., Sforzi A.;
RT "Genetic identification of wild and domestic cats (Felis silvestris), and their hybrids using Bayesian clustering methods.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ409135; CAC41048.1; -.
DR EMBL; AJ409140; CAC41063.1; -.
DR EMBL; AJ409142; CAC41069.1; -.
DR EMBL; AJ409144; CAC41075.1; -.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 8 AA; 951 MW; 262685BDC5A3733B CRC64;

Query Match 27.8%; Score 15; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;

Matches: 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WS 4
||
Db 3 WS 4

RESULT 15

Q94PX5 PRELIMINARY; PRT; 8 AA.
AC Q94PX5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT II (FRAGMENT).
GN COI.
OS Felis silvestris (Wild cat).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9683;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=66, 71, 75, 90, 1, AND 2;
RA Randi E., Pierpaoli M., Beaumont M., Ragni B., Sforzi A.;
RT "Genetic identification of wild and domestic cats (Felis silvestris),
RT and their hybrids using Bayesian clustering methods.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ409136; CAC41051.1; -
DR EMBL; AJ409137; CAC41054.1; -
DR EMBL; AJ409138; CAC41057.1; -
DR EMBL; AJ409139; CAC41060.1; -
DR EMBL; AJ409141; CAC41066.1; -
DR EMBL; AJ409143; CAC41072.1; -
KW Mitochondrion.
FT NON_TER
SQ SEQUENCE 8 AA; 951 MW; 262685BDC5A3733B CRC64;

Query Match 27.8%; Score 15; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WS 4
||
Db 3 WS 4

Search completed: October 10, 2002, 16:39:08
Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:10:16 ; Search time 27 seconds
(without alignments)
41.138 Million cell updates/sec

Title: US-09-848-834A-1
Perfect score: 54
Sequence: 1 XHWSYGLRPX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 135323

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	96.3	9	2 AAP10414	Luteinising Hormon
2	52	96.3	9	6 AAP50568	Sequence of gonado
3	52	96.3	9	20 AAW94891	LHRH peptide fragm
4	52	96.3	9	21 AAB15363	Human LHRH peptide
5	52	96.3	9	21 AAB08104	Amino acid sequenc
6	52	96.3	9	22 AAB90972	Luteinising hormon
7	52	96.3	9	22 AAB90979	Luteinising hormon
8	52	96.3	9	22 AAB59836	GnRH peptide, Pet
9	52	96.3	10	2 AAP10097	Sequence of lutein
10	52	96.3	10	2 AAP10411	Luteinising Hormon
11	52	96.3	10	2 AAP10416	Luteinising Hormon

12	52	96.3	10	3 AAP20277	Modified carboxy t
13	52	96.3	10	6 AAP50222	Gonadotrophin rele
14	52	96.3	10	7 AAP60127	Gonadoliberin anta
15	52	96.3	10	7 AAP61403	Gonadotrophin rele
16	52	96.3	10	7 AAP60576	Novel decapeptide
17	52	96.3	10	8 AAP70922	Luteinising hormon
18	52	96.3	10	10 AAP90630	Sequence of lutein
19	52	96.3	10	12 AAR15713	Peptide #1 with ho
20	52	96.3	10	13 AAR26819	LH releasing hormo
21	52	96.3	10	14 AAR33434	Therapeutic agent
22	52	96.3	10	15 AAR62689	LHRH hapten for at
23	52	96.3	10	16 AAR91197	LHRH peptide, Syn
24	52	96.3	10	16 AAR86845	Gonadotrophin rele
25	52	96.3	10	16 AAR75152	Gonadotrophin rele
26	52	96.3	10	17 AAW65201	Luteinising hormon
27	52	96.3	10	17 AAW65203	Luteinising hormon
28	52	96.3	10	18 AAW45642	Luteinising hormon
29	52	96.3	10	18 AAW22390	Gonadotrophin rele
30	52	96.3	10	18 AAW16390	Gonadotrophin rele
31	52	96.3	10	18 AAW04612	Luteinizing hormone
32	52	96.3	10	19 AAW79566	GnRH-1 polypeptide
33	52	96.3	10	19 AAW76381	Rat modified GnRH
34	52	96.3	10	19 AAW76373	Rat GnRH peptide.
35	52	96.3	10	19 AAW61541	Peptide hormone Gn
36	52	96.3	10	20 AAY50229	Neutrophil-activat
37	52	96.3	10	20 AAY31176	Ubiquitin fusion p
38	52	96.3	10	20 AAY31180	Ubiquitin fusion p
39	52	96.3	10	20 AAY31067	Non-crosslinked pr
40	52	96.3	10	20 AAY03864	Amino acid sequenc
41	52	96.3	10	20 AAY03856	LHRH peptide fragm
42	52	96.3	10	20 AAW94890	Luteinising hormon
43	52	96.3	10	20 AAW96765	Hormone domain of
44	52	96.3	10	20 AAW84278	Modified hormone d
45	52	96.3	10	20 AAW84286	

ALIGNMENTS

RESULT 1
AAP10414
ID AAP10414 standard; Protein; 9 AA.
XX
AC AAP10414;
XX
DT 17-DEC-1992 (first entry)
XX
DE Luteinising Hormone Releasing Hormone analogue #3.
XX
KW LHRH; Follicle Stimulating Factor; FSH; acne; hirsutism;
KW dysmennorhea; precocious puberty; endometriosis; prostate cancer;
KW benign prostate hypertrophy; mammary tumour.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /label= OTHER
FT /note= "pyroglutamic acid"
FT Modified-site 9 /note= "Pro-NH-(CH2)n-CH3 (n=0-2),
FT /note= "Pro-NH-(CH2)2-OH or protected by
FT pyrrolidino or morpholino gp."
XX
BE885308-A.
XX
PD 19-MAR-1981.
XX
PF 23-FEB-1983; 83BE-0468932.
XX
PR 21-SEP-1979; 79FR-0023545.
XX
PA (ROUS) ROUSSEL UCLAF.
XX
DR WPI; 1981-23409D/14 (23409D).

SQ Sequence 9 AA;
 Query Match 96.3%; Score 52; DB 6; Length 9;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRP 9
 |||||
 DB 2 HWSYGLRP 9
 |||||
 RESULT 3
 AAW94891
 ID AAW94891 standard; peptide; 9 AA.
 XX
 AC AAW94891;
 XX
 DT 11-MAY-1999 (first entry)
 XX
 DE LHRH peptide fragment.
 XX
 LHRH; immune response; luteinising hormone releasing hormone; DT;
 diphtheria toxoid; castrating; oestrus cycling; aggression; breast;
 KW sexual activity; organoleptic; livestock; cell growth; malignant;
 KW prostate; ovarian; oncofoetal; hyperplastic; pregnancy;
 KW endometriosis; inflammatory response.
 XX
 OS Homo sapiens.
 XX
 PN WO9902180-A1.
 XX
 PD 21-JAN-1999.
 XX
 PF 09-JUL-1998; 98WO-AU00532.
 XX
 PR 09-JUL-1997; 97AU-0007768.
 XX
 PA (CSLC-) CSL LTD.
 XX
 PI McNamara MK;
 XX
 DR WPI; 1999-120511/10.
 XX
 PT New immunogenic leutenising hormone releasing hormone compositions -
 PT comprise LHRH conjugated to diphtheria toxoid and adsorbed to an
 PT ionic polysaccharide, used to inhibit reproductive function in
 PT animals
 XX
 PS Example 3; Page 30; 4lpp; English.
 XX
 CC The invention relates immunogenic composition for eliciting an immune
 CC response to luteinising hormone releasing hormone (LHRH). The
 CC composition comprises a LHRH-diphtheria toxoid (DT) conjugate adsorbed to
 CC an ionic polysaccharide. The LHRH-DT compositions can be used for
 CC eliciting an immune response to LHRH, for castrating an animal, for
 CC regulating oestrus cycling in a female animal or for inhibiting
 CC characteristics induced by the sexual maturation of an animal, e.g.
 CC aggression or sexual activity. They can also be used for achieving
 CC production gains in livestock, e.g. reduction or elimination of unwanted
 CC organoleptic characteristics from the meat of livestock. They can also be
 CC used for inhibiting the growth of cells which are regulated directly or
 CC indirectly by LHRH, e.g. malignant breast cells, malignant prostate
 CC cells, malignant ovarian cells, malignant oncofoetal cells or
 CC hyperplastic cells. They can also be used for down-regulating the libido
 CC of an animal. They can also be used for inhibiting pregnancy, prostate
 CC enlargement, endometriosis or inflammatory responses. The LHRH
 CC compositions induce a more effective immune response against LHRH than
 CC the LHRH-carrier-adjvant compositions. The effective immune response
 CC against LHRH results in prevention of the release of the hormones LH and
 CC FSH from the anterior pituitary. Sequences AAW94890-93 are peptide
 CC derivatives of LHRH.
 XX
 SO Sequence 9 AA;

Query Match 96.3%; Score 52; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
DB 1 HWSYGLRP 8

RESULT 4

AAB15363
ID AAB15363 standard; peptide; 9 AA.

AC AAB15363;

DT 17-JAN-2001 (first entry)

DE Human LHRH peptide SEQ ID NO: 2.

Human: LHRH; GnRH; luteinising hormone releasing hormone;
gonadotrophin releasing hormone; fertility control; cancer;
endometriosis; prostate enlargement.

OS Homo sapiens.

XX WO200041720-A1.

XX 20-JUL-2000.

XX 24-DEC-1999; 99WO-AU01167.

XX 08-JAN-1999; 99AU-0008073.

XX (CSLC-) CSL LTD.

XX Walker J;

XX WPI; 2000-475954/41.

Adjuvant composition for manufacturing an immunogenic composition that
can elicit an immune response in an animal, comprises an ionic
polysaccharide component and a saponin component that is an
immunostimulating complex -

PS Disclosure; Page 50; 53pp; English.

The present sequence is a peptide fragment of human luteinising hormone
releasing hormone (also known as LHRH, GnRH and gonadotrophin releasing
hormone). It was used to demonstrate the novel adjuvant of the invention,
which has lower reactivity than previous compositions. Vaccination of
humans and animals against LHRH can be used as a method of fertility
control, as well as enabling the control and treatment of disorders of
the reproductive organs, such as testicular, breast, prostate and ovarian
cancers, prostate enlargement and endometriosis. The composition of the
invention contains an anionic macromolecule and a saponin component, the
latter of which is an immunostimulant, and it can also be used with other
immunogens including soluble protein antigens, peptide haptens conjugated
to a carrier protein and whole viruses.

XX Sequence 9 AA;

Query Match 96.3%; Score 52; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
DB 1 HWSYGLRP 8

RESULT 5

AAB08104

ID AAB08104 standard; peptide; 9 AA.

XX AAB08104;

DT 04-DEC-2000 (first entry)

DE Amino acid sequence of truncated luteinising hormone releasing hormone.
T helper cell epitope; CDV; immune response; canine vaccine;
luteinising hormone releasing hormone; LHRH.

OS Canis sp.

XX WO200046390-A1.

XX 10-AUG-2000.

XX 07-FEB-2000; 2000WO-AU00070.

XX 05-FEB-1999; 99AU-0008533.

XX 04-AUG-1999; 99AU-0002013.

XX (UYME) UNIV MELBOURNE.

XX (CSLC-) CSL LTD.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

XX Jackson DC, Souravi G, Walker J;

XX WPI; 2000-532904/48.

XX Novel T helper cell epitopes derived from canine distemper virus useful
for preparation of canine vaccines -

XX Example 3; Page 21; 54pp; English.

The present sequence represents luteinising hormone releasing hormone
(LHRH). It is used in vaccines with T helper cell epitopes
comprising these T cell helper epitopes are useful for inducing an
immune response in an animal. The epitopes are useful as components
of animal, in particular, canine vaccines, either simply as synthetic
peptide based vaccines and as additions to vaccines containing more
complex antigens.

XX Sequence 9 AA;

Query Match 96.3%; Score 52; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
DB 1 HWSYGLRP 8

RESULT 6

AAB90972

ID AAB90972 standard; Peptide; 9 AA.

XX AAB90972;

DT 22-JUN-2001 (first entry)

DE Luteinising hormone releasing hormone (LH-RH) related peptide SEQ ID:146.

Protection; endogenous therapeutic peptide; peptidase; conjugation;
blood component; modification; succinimidy; maleimido group; amino;
hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

OS Synthetic.

```
XX PN WO200069900-A2.
XX XX
XX PD 23-NOV-2000.
XX PF 17-MAY-2000; 2000WO-US13576.
XX XX
XX PR 17-MAY-1999; 99US-0134406.
XX PR 10-SEP-1999; 99US-0153406.
XX PR 15-OCT-1999; 99US-0159783.
XX PA (CONJ-) CONJUCHEM INC.
XX XX
XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX XX
XX DR WPI; 2001-112059/12.
XX XX
XX PT Modifying and attaching therapeutic peptides to albumin prevents
XX PT peptidase degradation, useful for increasing length of in vivo activity
XX XX
XX PS Disclosure; Page 238; 733pp; English.
XX CC The present invention describes a modified therapeutic peptide (I)
XX CC comprising a therapeutically active amino acid region (III) and a
XX CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
XX CC a less therapeutically active amino acid region (IV), which covalently
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
XX CC factors and neurotransmitters, to protect them from peptidase activity
XX CC in vivo for the treatment of various disorders. Endogenous therapeutic
XX CC peptides are not suitable as drug candidates as they require frequent
XX CC administration due to rapid degradation by peptidases in the body.
XX CC Modifying and attaching therapeutic peptides to albumin prevents or
XX CC reduces the action of peptidases to increase length of activity (half
XX CC life) and specificity as bonding to large molecules decreases
XX CC intracellular uptake and interference with physiological processes.
XX CC AAB90829 to AAB92441 represent peptides which can be used in the
XX CC exemplification of the present invention.
XX SQ Sequence 9 AA;

Query Match 96.3%; Score 52; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db |||||
2 HWSYGLRP 9

RESULT 7
AAB90979
ID AAB90979 standard; Peptide; 9 AA.
XX AC AAB90979;
XX XX
XX DT 22-JUN-2001 (first entry)
XX DE Luteinising hormone releasing hormone (LH-RH) related peptide SEQ ID:153.
XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
XX KW blood component; modification; succinimidyl; maleimido group; amino;
XX KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX XX
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200069900-A2.
XX XX
XX PD 23-NOV-2000.
XX PF 17-MAY-2000; 2000WO-US13576.
XX XX
XX PR 17-MAY-1999; 99US-0134406.
XX PR 10-SEP-1999; 99US-0153406.
XX PR 15-OCT-1999; 99US-0159783.
XX PA (CONJ-) CONJUCHEM INC.
XX XX
XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX XX
XX DR WPI; 2001-112059/12.
XX XX
XX PT Modifying and attaching therapeutic peptides to albumin prevents
XX PT peptidase degradation, useful for increasing length of in vivo activity
XX XX
XX PS Disclosure; Page 238; 733pp; English.
XX CC The present invention describes a modified therapeutic peptide (I)
XX CC comprising a therapeutically active amino acid region (III) and a
XX CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
XX CC a less therapeutically active amino acid region (IV), which covalently
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
XX CC factors and neurotransmitters, to protect them from peptidase activity
XX CC in vivo for the treatment of various disorders. Endogenous therapeutic
XX CC peptides are not suitable as drug candidates as they require frequent
XX CC administration due to rapid degradation by peptidases in the body.
XX CC Modifying and attaching therapeutic peptides to albumin prevents or
XX CC reduces the action of peptidases to increase length of activity (half
XX CC life) and specificity as bonding to large molecules decreases
XX CC intracellular uptake and interference with physiological processes.
XX CC AAB90829 to AAB92441 represent peptides which can be used in the
XX CC exemplification of the present invention.
XX SQ Sequence 9 AA;

Query Match 96.3%; Score 52; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db |||||
2 HWSYGLRP 9

RESULT 8
AAB59836
ID AAB59836 standard; Peptide; 9 AA.
XX AC AAB59836;
XX XX
XX DT 26-MAR-2001 (first entry)
XX DE GnRH peptide.
XX XX
XX KW GnRH-III; autoimmune disease; transplant rejection; retroviral disease;
XX KW graft-versus-host-disease; lymphoproliferative disease;
XX KW gonadotropin-releasing hormone.
XX XX
XX OS Petromyzon marinus.
XX XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "Linked to Glucagon-like peptide"
XX FT
XX XX
XX PN WO200074724-A2.
XX XX
XX PD 14-DEC-2000.
XX XX
XX PF 05-JUN-2000; 2000WO-GB02014.
XX XX
```



```
XX SQ Sequence 10 AA;
Query Match 96.3%; Score 52; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
   |||||
Db 2 HWSYGLRP 9

RESULT 11
AAP10416
ID AAP10416 standard; peptide; 10 AA.
XX
AC AAP10416;
XX
DT 17-DEC-1992 (first entry)
XX
DE Luteinising Hormone Releasing Hormone analogue #5.
XX
KW LHRH; Follicle Stimulating Factor; FSH; acne; hirsutism;
KW dysmenorrhea; precocious puberty; endometriosis; prostate cancer;
KW benign prostate hypertrophy; mammary tumour.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "pyroglutamic acid"
FT Modified-site 7
FT /label= OTHER
FT /note= "N-alpha-methyl-Leu"
FT Modified-site 10
FT /note= "amidated or absent, in which case Pro(9)
FT is Pro-NH-C2H5"
XX
PN BE885308-A.
XX
PD 19-MAR-1981.
XX
PF 23-FEB-1983; 83BE-0468932.
XX
PR 21-SEP-1979; 79FR-0023545.
XX
PA (ROUS ) ROUSSEL UCLAF.
XX
DR WPI; 1981-23409D/14 (23409D).
XX
PT LH-RH, liberating factor for LH and FSH, and its agonists compsn.
PT - used to treat prostate adenocarcinoma, benign hypertrophy of
PT the prostate, hirsutism, acne, etc.
XX
PS Claim 1(f); Page 16; 27pp; French.
XX
CC A composition is claimed containing LHRH or its analogues. The
CC composition is used to treat prostate adenocarcinoma, benign
CC hypertrophy of the prostate, endometriosis, dysmenorrhea, hirsutism,
CC hormone-dependent mammary tumours, for treatment or prevention of
CC precocious puberty, delaying the onset of puberty and for treating
CC acne. The compositions may also contain antiandrogens.
CC See AAP10411-P10418.
XX
SQ Sequence 10 AA;
Query Match 96.3%; Score 52; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
   |||||
Db 2 HWSYGLRP 9

RESULT 12
AAP20277
ID AAP20277 standard; Protein; 10 AA.
XX
AC AAP20277;
XX
DT 30-NOV-1992 (first entry)
XX
DE Modified carboxy terminal peptide 2.
XX
KW Medicament; pituitary function; hypothalamic releasing factors;
KW enkephalin.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 10
FT /label= Modified_gly_with_terminal_CHN2_or_CH2X
FT /note= "X= Cl, Br, I"
XX
PN US4305872-A.
XX
PD 15-DEC-1981.
XX
PF 19-OCT-1979; 79US-0086417.
XX
PR 19-OCT-1979; 79US-0086417.
XX
PA (WING/) WINGROVE K.
XX
PI Johnston RB, Balk JI, Pelton JT;
XX
DR WPI; 1982-01722E/01 (01722E).
XX
PT Diazo- and halo-methyl ketone derivs. of polypeptide(s) - useful
PT as hormone or opiate antagonists or agonists
XX
PS Claim 2; Page 25; 26pp; English.
XX
CC The sequences given in AAP20276-80 are biologically active derivatives
CC of carboxy-terminal polypeptides which have the OH group of the
CC terminal carboxy group replaced by CHN2 or CH2X (where X = Cl, Br or
CC I). These peptides are agonists or antagonists to the polypeptides
CC from which they are derived. They can be used as medicaments eg.
CC for regulating pituitary function, or esp. as research tools for
CC investigating the action of the polypeptides in biological systems.
CC The polypeptides used to obtain the derivatives were selected from
CC acid free forms of hypothalamic releasing factors, enkephalins and
CC biologically active polypeptide fragments of these.
XX
SQ Sequence 10 AA;
Query Match 96.3%; Score 52; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
   |||||
Db 2 HWSYGLRP 9

RESULT 13
AAP50222
ID AAP50222 standard; Protein; 10 AA.
XX
AC AAP50222;
XX
DT 20-JAN-1992 (first entry)
XX
DE Gonadotrophin release stimulating hormone.
XX
KW GnRH; LH-RH; LRF; gonadotrophins; steroids; contraceptive.
```

XX OS Synthetic.
 XX PN EPI43573-A.
 XX PD 05-JUN-1985.
 XX XX 05-NOV-1984; 84EP-0307625.
 XX PR 29-NOV-1983; 83US-0556148.
 XX PR 30-AUG-1985; 85US-0771517.
 XX PA (SALK) SALK INST FOR BIOL STUD.
 XX PI Roeske RW, Rivier JE, Vale WW;
 XX XX WPI; 1985-136434/23.
 XX DR New GnRH antagonist peptide(s) - useful as inhibitors of
 PT gonadotropin(s) and/or steroid(s) for contraceptive use.
 XX PS Disclosure; Page 1; 20pp; English.
 XX XX The claimed peptide antagonists inhibit the release of gonadotrophins
 CC and/or steroids. They are antagonistic to GnRH, inhibit ovulation, and
 CC may cause resorption of a fertilised egg if administered shortly after
 CC absorption. The peptides also have utility in male contraception, and
 CC in treatment of precocious puberty, hormone dependent neoplasia,
 CC dysmenorrhoea and endometriosis.
 XX XX Sequence 10 AA;
 SQ Query Match 96.3%; Score 52; DB 6; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRP 9
 Db | | | | | | | |
 2 HWSYGLRP 9
 RESULT 14
 AAP60127
 ID AAP60127 standard; Peptide; 10 AA.
 XX AC AAP60127;
 XX DT 12-JUN-1991 (first entry)
 XX DE Gonadoliberein antagonist.
 XX KW Gonadoliberein antagonist; contraceptive; antitumor.
 XX PN EP201260-A.
 XX PD 12-NOV-1986.
 XX XX 28-APR-1986; 86EP-0303210.
 XX PR 09-MAY-1985; 85US-0732531.
 XX XX (SALK) SALK INST FOR BIOL STUD.
 XX PA Rivier JEF, Varga JI, Hagler AT, Struthers RS, Perrin MH;
 XX PI Rivier CL, Vale WW;
 XX XX WPI; 1986-299774/46.
 XX DR New peptide gonadotropin releasing hormone antagonists - useful
 PT esp. as contraceptives, for treating early puberty,
 PT hormone-dependent neoplasms etc.
 XX PS Disclosure; Page 1; 33pp; English.

XX CC The decapeptide encodes a gonadoliberein antagonist, which may be
 CC used as a male contraceptive and as an antitumour (against steroid-
 CC dependent tumours).
 XX SQ Sequence 10 AA;
 Query Match 96.3%; Score 52; DB 7; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRP 9
 Db | | | | | | | |
 2 HWSYGLRP 9
 RESULT 15
 AAP61403
 ID AAP61403 standard; protein; 10 AA.
 XX AC AAP61403;
 XX DT 04-AUG-1991 (first entry)
 XX DE Gonadotropin releasing hormone.
 XX KW Gonadotropin releasing hormone; analogue; peptide synthesis;
 KW ovulation; veterinary medicine; fertility;
 XX PN DD232500-A.
 XX XX 29-JAN-1986.
 XX PF 08-MAY-1984; 84DD-0262804.
 XX PR 08-MAY-1984; 84DD-0262804.
 XX PA (DEAK) AKAD WISSENSCHAFT DDR.
 XX PI Kaufmann KD, Dolling R, Handel L;
 XX XX WPI; 1986-137868/22.
 XX DR Prepn. of gonadotropin liberating hormone and analogues - by
 PT multistage rapid peptide synthesis in soln. without isolating
 PT intermediates
 XX PS Disclosure; page 7; 8pp; german.
 XX CC The gonadotropin releasing hormone and its analogues are prepd. by a
 CC new multistage rapid peptide synthesis method in soln., where the
 CC intermediates are not isolated. The process is rapid and gives very
 CC pure peptide quickly and using little equipment. The peptide can be
 CC used in veterinary medicine to synchronise ovulation in large animal
 CC herds, and in human medicine in the treatment of fertility disorders.
 XX SQ Sequence 10 AA;
 Query Match 96.3%; Score 52; DB 7; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRP 9
 Db | | | | | | | |
 2 HWSYGLRP 9
 Search completed: October 10, 2002, 16:37:45
 Job time : 28 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: October 10, 2002, 16:37:17 ; Search time 16 Seconds
(without alignments)
15.266 Million cell updates/sec

Title: US-09-848-834A-1
Perfect score: 54
Sequence: 1 XHWSYGLRPX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 70601

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	96.3	10	1	US-07-714-540-9
2	52	96.3	10	1	US-07-983-111-1
3	52	96.3	10	1	US-07-690-983D-1
4	52	96.3	10	1	US-07-690-983D-2
5	52	96.3	10	1	US-07-690-983D-6
6	52	96.3	10	1	US-07-690-983D-7
7	52	96.3	10	1	US-07-690-983D-8
8	52	96.3	10	1	US-07-690-983D-32
9	52	96.3	10	1	US-07-690-983D-37
10	52	96.3	10	1	US-08-103-022-1
11	52	96.3	10	1	US-07-897-680-1
12	52	96.3	10	1	US-08-184-935-6
13	52	96.3	10	1	US-08-343-883-1
14	52	96.3	10	1	US-08-000-931-5
15	52	96.3	10	1	US-08-428-488-22
16	52	96.3	10	1	US-08-341-219-11
17	52	96.3	10	1	US-08-453-588-2
18	52	96.3	10	1	US-08-453-588-4
19	52	96.3	10	1	US-08-453-588-6
20	52	96.3	10	1	US-08-453-588-8
21	52	96.3	10	1	US-08-453-588-10
22	52	96.3	10	1	US-08-453-588-12
23	52	96.3	10	1	US-08-453-588-14
24	52	96.3	10	1	US-08-453-588-16
25	52	96.3	10	1	US-08-453-588-19
26	52	96.3	10	1	US-08-453-588-22
27	52	96.3	10	1	US-08-188-223-1

28	52	96.3	10	1	US-08-188-223-3	Sequence 3, Appli
29	52	96.3	10	1	US-08-188-223-8	Sequence 8, Appli
30	52	96.3	10	1	US-08-406-935-5	Sequence 5, Appli
31	52	96.3	10	1	US-08-591-917-1	Sequence 1, Appli
32	52	96.3	10	1	US-08-387-156-2	Sequence 2, Appli
33	52	96.3	10	1	US-08-474-555-1	Sequence 1, Appli
34	52	96.3	10	1	US-08-446-692-1	Sequence 1, Appli
35	52	96.3	10	1	US-08-242-678D-1	Sequence 1, Appli
36	52	96.3	10	2	US-08-796-598-6	Sequence 2, Appli
37	52	96.3	10	2	US-08-694-865-2	Sequence 6, Appli
38	52	96.3	10	2	US-08-694-865-18	Sequence 18, Appli
39	52	96.3	10	2	US-08-488-351A-1	Sequence 1, Appli
40	52	96.3	10	2	US-08-480-494B-1	Sequence 1, Appli
41	52	96.3	10	2	US-08-447-175A-6	Sequence 6, Appli
42	52	96.3	10	2	US-08-747-137-116	Sequence 116, App
43	52	96.3	10	2	US-08-878-748-2	Sequence 2, Appli
44	52	96.3	10	3	US-08-521-079-2	Sequence 2, Appli
45	52	96.3	10	3	US-08-521-079-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-07-714-540-9
; Sequence 9, Application US/07714540
; Patent No. 5262521
; GENERAL INFORMATION:
; APPLICANT: Almquist, Ronald G.
; TITLE OF INVENTION: ISOLATED ATRIAL PEPTIDE-DEGRADING
; TITLE OF INVENTION: ISOLATED ATRIAL PEPTIDE-DEGRADING
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Irell & Manella
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/714,540
; FILING DATE: 19910607
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Dianne E.
; REGISTRATION NUMBER: 31,292
; REFERENCE/DOCKET NUMBER: 8500-0135.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-714-540-9

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
|||||||

DB 2 HWSYGLRP 9

RESULT 2
US-07-983-111-1
; Sequence 1, Application US/07983111
; Patent No. 5284657
; GENERAL INFORMATION:
; APPLICANT: Fu Lu, Mou-Ying
; APPLICANT: Relland, Thomas L.
; TITLE OF INVENTION: "Compositions and Methods for the
; TITLE OF INVENTION: Sublingual or Buccal Administration of Therapeutic Agents"
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman, Jr., Dept. 377-AP6D
; STREET: One Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/983,111
; FILING DATE: 30-NOV-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/750,843
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Janssen, Jerry F.
; REGISTRATION NUMBER: 29,175
; REFERENCE/DOCKET NUMBER: 4848.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 938-7742
; TELEFAX: (708) 937-9556
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa at position 1 is
; OTHER INFORMATION: 5-oxo-proline"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "Xaa at position 10 is
; OTHER INFORMATION: glycineamide"
US-07-983-111-1
Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRP 9
Db 2 HWSYGLRP 9
RESULT 3
US-07-690-983D-1
; Sequence 1, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Represents pyroglutamic
; OTHER INFORMATION: acid"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "Represents glycineamide"
US-07-690-983D-1
Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRP 9
Db 2 HWSYGLRP 9
RESULT 4
US-07-690-983D-2
; Sequence 2, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/690,983D
;; FILING DATE: 25-JUN-1991
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/AU90/00373
;; FILING DATE: 24-AUG-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
US-07-690-983D-2

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | | | | |
DB 2 HWSYGLRP 9

RESULT 5
US-07-690-983D-6
;; Sequence 6, Application US/07690983D
;; Patent No. 5403586
;; GENERAL INFORMATION:
;; APPLICANT: RUSSELL-JONES, Gregory J.
;; APPLICANT: STEWART, Andrew G.
;; APPLICANT: TSONIS, Con G.
;; TITLE OF INVENTION: FUSION PROTEINS
;; NUMBER OF SEQUENCES: 47
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 3000 K Street, N.W.
;; CITY: Washington, D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/690,983D
;; FILING DATE: 25-JUN-1991
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/AU90/00373
;; FILING DATE: 24-AUG-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown

;; MOLECULE TYPE: protein
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 10
;; OTHER INFORMATION: /note= "Represents glycinamide"
US-07-690-983D-6

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | | | | |
DB 2 HWSYGLRP 9

RESULT 6
US-07-690-983D-7
;; Sequence 7, Application US/07690983D
;; Patent No. 5403586
;; GENERAL INFORMATION:
;; APPLICANT: RUSSELL-JONES, Gregory J.
;; APPLICANT: STEWART, Andrew G.
;; APPLICANT: TSONIS, Con G.
;; TITLE OF INVENTION: FUSION PROTEINS
;; NUMBER OF SEQUENCES: 47
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 3000 K Street, N.W.
;; CITY: Washington, D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/690,983D
;; FILING DATE: 25-JUN-1991
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/AU90/00373
;; FILING DATE: 24-AUG-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 10
;; OTHER INFORMATION: /note= "Represents glycinamide"
US-07-690-983D-7

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | | | | |
DB 2 HWSYGLRP 9

RESULT 7

US-07-690-983D-8
; Sequence 8, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690.983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "Represents glycylamide"
US-07-690-983D-8

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 8
US-07-690-983D-32
; Sequence 32, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690.983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-690-983D-32

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 9
US-07-690-983D-37
; Sequence 37, Application US/07690983D-
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690.983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-690-983D-37

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | | | | |
Db 3 HWSYGLRP 10

RESULT 10

US-08-103-022-1
; Sequence 1, Application US/08103022
; Patent No. 5413990
; GENERAL INFORMATION:
; APPLICANT: Havig, Fortuna
; APPLICANT: Fitzpatrick, Timothy D.
; APPLICANT: Swenson, Rolf E.
; APPLICANT: Nichols, Charles J.
; APPLICANT: Mort, Nicholas A.
; TITLE OF INVENTION: N-terminus Modified Analogs of LHRH
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edward H. Gorman, Jr., Dept. 377
; STREET: Abbott Laboratories, One Abbott Park Road
; CITY: No. 5413990th Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103/022
FILING DATE: 05-OCT-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 5389.US.O1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 938-2623
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1

OTHER INFORMATION: /note= "xaa at position 1 is a
5-oxo-prolyl aminoacyl residue."
US-08-103-022-1
Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | | | | |
Db 2 HWSYGLRP 9

RESULT 11

US-07-897-680-1

; Sequence 1, Application US/07897680
; Patent No. 5446025
; GENERAL INFORMATION:
; APPLICANT: Fu Lu, Mou-Ying
; APPLICANT: Subba Rao, Gowdahallin N.
; APPLICANT: Lee, Dennis Y.
; TITLE OF INVENTION: Formulations and Method for the
; TITLE OF INVENTION: Percutaneous Administration of Leuprolide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dept. 377 Abbott Laboratories
; STREET: One Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/897,680
FILING DATE: 19920612
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 5165.US.O1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 937-9556
TELEFAX: (708) 938-7742
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "XAA at position 1 is a
pyro-glutamyl residue"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note= "XAA at position 10 is a
glycyl-amide residue"
US-07-897-680-1

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
| | | | | | | |
Db 2 HWSYGLRP 9

RESULT 12

US-08-184-935-6
; Sequence 6, Application US/08184935
; Patent No. 5476770
; GENERAL INFORMATION:
; APPLICANT: PRADELLES, PHILIPPE
; TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN
; TITLE OF INVENTION: OR HAPTEN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400

; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,935
; FILING DATE: 24-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5476770man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 846-286-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "C-terminal amide"
; US-08-184-935-6

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 13
US-08-343-883-1
; Sequence 1, Application US/08343883
; Patent No. 5573767
; GENERAL INFORMATION:
; APPLICANT: Dufour, Raymond J.
; APPLICANT: Roulet, Claude J.M.
; APPLICANT: Chouvet, Claire D.
; APPLICANT: Bonneau, Michel B.
; TITLE OF INVENTION: Method for improving the organoleptic
; TITLE OF INVENTION: qualities of the meat from uncastrated male domestic
; TITLE OF INVENTION: animals, vaccines which are usable in this method, new
; TITLE OF INVENTION: peptide, in particular for producing these vaccines...
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Larson and Taylor
; STREET: 727 Twenty-Third Street, South
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,883
; FILING DATE: 17-NOV-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/946,495
; FILING DATE: 09-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9102513
; FILING DATE: 01-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 9115289
; FILING DATE: 10-DEC-1991
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 10
; OTHER INFORMATION: /label= NH2
; OTHER INFORMATION: /note= "amidated glycine"
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /label= pyro
; OTHER INFORMATION: /note= "pyroglutamic acid"
; PUBLICATION INFORMATION:
; AUTHORS: Matsuo, H.
; AUTHORS: Baba, Y.
; AUTHORS: G. Nair, R. M.
; AUTHORS: Arimura, A. V.
; AUTHORS: Schally, A. V.
; TITLE: Structure of the porcine LH- and
; TITLE: FSH-releasing hormone. I. The proposed amino acid
; TITLE: sequence.
; JOURNAL: Biochem. Biophys. Res. Commun.
; VOLUME: 43
; ISSUE: 6
; PAGES: 1334-1339
; DATE: 1971
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 10
; US-08-343-883-1

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 14
US-08-000-931-5
; Sequence 5, Application US/08000931
; Patent No. 5578477
; GENERAL INFORMATION:
; APPLICANT: Tamanoi Dr., Fuyuhiko
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF
; TITLE OF INVENTION: INHIBITORS OF PROTEIN FARNESYLTRANSFERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/000,931
; FILING DATE: 05-JAN-1994

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 64098/102/ARDE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-000-931-5

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

RESULT 15
US-08-428-488-22
Sequence 22, Application US/08428488
Patent No. 5624894
GENERAL INFORMATION:
APPLICANT: BODOR, Nicholas S.
TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE
TITLE OF INVENTION: PEPTIDES BY SEQUENTIAL METABOLISM
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,488
FILING DATE: 27-APR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Mary Katherine
REGISTRATION NUMBER: 26,254
REFERENCE/DOCKET NUMBER: 028724-087
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Position 1 = p-Glu."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10

OTHER INFORMATION: /note= "Position 10 = Gly-NH2."
US-08-438-488-22

Query Match 96.3%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRP 9
Db 2 HWSYGLRP 9

Search completed: October 10, 2002, 16:40:21
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:45:57 ; Search time 12.5 Seconds
(without alignments)
238.302 Million cell updates/sec

Title: US-09-848-834A-9

Perfect score: 160

Sequence: 1 KLLSEIKGVIVHRLEGVEGSLHWSYGLRXP 31

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 6607

Minimum DB seq length: 0
Maximum DB seq length: 31

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	32.5	10	1 RHPGG	gonadoliberin - pig
2	52	32.5	10	1 RHSHG	gonadoliberin - sh
3	48	30.0	10	1 RHAQ1	gonadoliberin I -
4	39	24.4	10	2 A21114	gonadoliberin - ch
5	38	23.8	21	2 A60225	pyruvate dehydroge
6	34	21.2	10	1 RHAQ2	gonadoliberin II -
7	34	21.2	10	1 A61126	gonadoliberin I -
8	34	21.2	10	2 A46030	gonadoliberin II -
9	34	21.2	10	2 B46030	ig heavy chain (cl
10	34	21.2	29	2 S10061	gonadoliberin I -
11	32	20.0	25	2 S29284	hydrogenase (EC 1.
12	31	19.4	10	2 A49187	gonadotropin-rela
13	30.5	19.1	18	2 S29379	sorbitol dehydroge
14	30	18.8	18	4 I39461	anti-angiotensin,
15	30	18.8	20	2 PS0188	superoxide dismuta
16	30	18.8	21	2 D42762	multicatalytic end
17	29	18.1	30	2 A49955	protein-tyrosine k
18	28.5	17.8	24	2 S10664	ribosomal protein
19	28	17.5	16	2 S57517	T cell receptor be
20	28	17.5	21	2 T11806	ribosomal protein
21	28	17.5	22	2 P00070	T-cell receptor be
22	28	17.5	25	2 B36934	orf3 3' of mada -
23	28	17.5	28	2 T06925	hypothetical prote
24	27.5	17.2	19	2 P50186	superoxide dismuta
25	27.5	17.2	27	2 B32112	R15 beta peptide -
26	27	16.9	14	2 PA0109	porin por 1B - Ara
27	27	16.9	14	2 PA0045	porin por1 - Arabi
28	27	16.9	17	2 S69164	ferredoxin al - Ja
29	27	16.9	20	2 S21244	H+-transporting AT

30 27 16.9 22 2 S45058 coat protein - tur
31 27 16.9 26 2 S27244 formylmethanofuran
32 27 16.9 27 2 A45140 fatty-acyl-ethyl-e
33 27 16.9 30 2 B29164 cartilage proteogl
34 26 16.2 17 2 I51203 myosin heavy chain
35 26 16.2 22 2 PH1325 Ig heavy chain DJ
36 26 16.2 22 2 I59594 monoamine oxidase
37 26 16.2 27 2 PNO583 tyrosine 3-monooxy
38 26 16.2 27 2 PNO584 tyrosine 3-monooxy
39 26 16.2 27 2 PNO585 tyrosine 3-monooxy
40 26 16.2 29 2 PS0134 H-2 class I histoc
41 26 16.2 30 2 S21195 spectrin beta chai
42 26 16.2 30 2 A31836 17k antigen - Rick
43 26 16.2 31 2 D95079 hypothetical prote
44 26 16.2 31 2 B82138 hypothetical prote
45 25.5 15.9 19 2 S19532 globin - polychaet

ALIGNMENTS

RESULT 1

RHPGG
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A:Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117344
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:10/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 32.5%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.29;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30

DB 2 HWSYGLRP 9

RESULT 2

RHSHG

gonadoliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997

C:Accession: A93780; A01411

R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972

A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa

A:Reference number: A93780; MUID:72094314

A:Accession: A93780

A:Molecule type: protein

```
A:Residues: 1-10 <BUR>
A:Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match      32.5%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30
DB 2 HWSYGLRP 9
|||||

RESULT 3
RHAQI
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson
Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains of
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match      30.0%; Score 48; DB 1; Length 10;
Best Local Similarity 87.5%; Pred. No. 1.1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30
DB 2 HWSYGLRP 9
|||||

RESULT 4
A21114
gonadoliberin - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 18-Jun-1993
C:Accession: A21114
R:Sherwood, N.; Eiden, L.; Brownstein, M.; Spiess, J.; Rivier, J.; Vale, W.
Proc. Natl. Acad. Sci. U.S.A. 80, 2794-2798, 1983
A:Title: Characterization of a teleost gonadotropin-releasing hormone.
A:Reference number: A21114; MUID:83195140
A:Accession: A21114
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SHE>

Query Match      24.4%; Score 39; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30
DB 2 HWSYGLRP 9
|||||

RESULT 5
A60225
pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) alpha chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
```

```
C:Date: 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 15-Oct-1999
C:Accession: A60225
R:Lawson, R.; Aitken, A.; Yeaman, S.J.
Biochem. Soc. Trans. 11, 298-299, 1983
A:Title: Primary sequence of the N-terminal region of the alpha-subunit of pyruvate
A:Reference number: A60225
A:Accession: A60225
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-21 <LAW>
C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphatase
C:Keywords: oxidoreductase

Query Match      23.8%; Score 38; DB 2; Length 21;
Best Local Similarity 62.5%; Pred. No. 62;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 5 EIKGVIVHRLEGVEGP 20
DB 8 EIKKCDLHRLE--EGP 21
|||||

RESULT 6
RHAQI
gonadoliberin II - American alligator
N:Alternate names: gonadotropin-releasing hormone II
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: B60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swanson
Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains
A:Reference number: A60066; MUID:91352338
A:Accession: B60066
A:Molecule type: protein
A:Residues: 1-10 <LOW>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match      21.2%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30
DB 2 HWSYGLRP 9
|||||

RESULT 7
A61126
gonadoliberin - spotted ratfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Hydrolyagus colliei (spotted ratfish)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 18-Mar-1997
C:Accession: A61126
R:Lovejoy, D.A.; Sherwood, N.M.; Fischer, W.H.; Jackson, B.C.; Rivier, J.E.; Lee, T.
Gen. Comp. Endocrinol. 82, 152-161, 1991
A:Title: Primary structure of gonadotropin-releasing hormone from the brain of a ho
A:Reference number: A61126; MUID:91340067
A:Accession: A61126
A:Molecule type: protein
A:Residues: 1-10 <LOW>
A:Experimental source: brain
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; brain; hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match      21.2%; Score 34; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```


QY 23 HWSYGLRP 30
|||:| |
Db 2 HWSHGWP 9

RESULT 8

A46030
gonadoliberin I - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: A46030
R:Lovejoy, D.A.; Fischer, W.H.; Nganvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: A46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 21.2%; Score 34; DB 2; Length 10;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30
|||:| |
Db 2 HWSHGWP 9

RESULT 9

B46030
gonadoliberin II - spiny dogfish
N:Alternate names: gonadotropin-releasing hormone
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Dec-1998
C:Accession: B46030
R:Lovejoy, D.A.; Fischer, W.H.; Nganvongchon, S.; Craig, A.G.; Nahorniak, C.S.; Peter, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6373-6377, 1992
A:Title: Distinct sequence of gonadotropin-releasing hormone (GnRH) in dogfish brain pro
A:Reference number: A46030; MUID:92335300
A:Accession: B46030
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Keywords: hormone; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 21.2%; Score 34; DB 2; Length 10;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30
|||:| |
Db 2 HWSHGWP 9

RESULT 10

S10061
Ig heavy chain (clone J12) - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 15-Oct-1999
C:Accession: S10061; E49021
R:Schwager, J.; Grossberger, D.; du Pasquier, L.
EMBO J. 7, 2409-2415, 1988
A:Title: Organization and rearrangement of immunoglobulin M genes in the amphibian Xenop
A:Reference number: S01156; MUID:89052653
A:Accession: S10061
A:Molecule type: mRNA
A:Residues: 1-29 <SCH>

A:Cross-references: EMBL:X14925; NID:964844; PIDN:CAA33052.1; PID:9930274
A:Note: the authors translated the codon AAG for residue 5 as Leu and GAG for residue
R:Haire, R.N.; Anemiya, C.T.; Suzuki, D.; Litman, G.W.
J. Exp. Med. 171, 1721-1737, 1990
A:Title: Eleven distinct V-H gene families and additional patterns of sequence variat
A:Reference number: A47624; MUID:90237760
A:Accession: E49021
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 10-24 <HAI>
A:Note: J8 region
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-24/Domain: V-D-J region (fragment) #status predicted <VRE>
F:25-29/Domain: C region (mu chain) (fragment) #status predicted <CRE>

Query Match 21.2%; Score 34; DB 2; Length 29;
Best Local Similarity 46.7%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 13 RUEGVGEPGLHWSYG 27
|:|:| | | |
Db 3 RYKGVRYFEHNGQG 17

RESULT 11
S29284
hydrogenase (EC 1.18.99.1) (Fe) small chain - Desulfovibrio desulfuricans (fragment)

N:Alternate names: Fe hydrogenase beta chain
C:Species: Desulfovibrio desulfuricans
C:Date: 19-Mar-1997 #sequence_revision 24-Jul-1998 #text_change 24-Jul-1998
C:Accession: S29284
R:Hatchikian, E.C.; Forget, N.; Fernandez, V.M.; Williams, R.; Cammack, R.
Eur. J. Biochem. 209, 357-365, 1992
A:Title: Further characterization of the [Fe]-hydrogenase from Desulfovibrio desulfur
A:Reference number: S29283; MUID:93011148
A:Accession: S29284
A:Molecule type: protein
A:Residues: 1-25 <HAT>

A:Experimental source: strain ATCC 7757
C:Comment: This hydrogenase complex contains three [4Fe-4S] iron-sulfur clusters.
C:Complex: heterodimer; large (alpha) and small (beta) chain
C:Function:
A:Description: involved in production or consumption of molecular hydrogen coupled to
A:Note: may be involved in hydrogen uptake for reduction of sulfate to hydrogen sulfi
C:Superfamily: hydrogenase (Fe) small chain
C:Keywords: heterodimer; hydrogen metabolism; iron-sulfur protein; oxidoreductase; pe

Query Match 20.0%; Score 32; DB 2; Length 25;
Best Local Similarity 35.3%; Pred. No. 5.3e+02;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 LSEIKGVIVHRLGVEG 19
:|:|:|:|:| | | |
Db 2 VKQIKDYMLDRINGVYG 18

RESULT 12

A49187
gonadotropin-releasing hormone III - sea lamprey
C:Species: Petromyzon marinus (sea lamprey)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A49187
R:Sower, S.A.; Chiang, Y.C.; Lovas, S.; Conlon, J.M.
Endocrinology 132, 1125-1131, 1993
A:Title: Primary structure and biological activity of a third gonadotropin-releasing
A:Reference number: A49187; MUID:93178316
A:Accession: A49187
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <SOW>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:126381)

Query Match 19.4%; Score 31; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30
|||: :|
Db 2 HWSHWKP 9

RESULT 13

S29379

sorbitol dehydrogenase - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 19-May-2000

C:Accession: S29379

R:Reiseren, H.; Sletten, K.; McKinley-McKee, J.S.

Eur. J. Biochem. 211, 861-869, 1993

A:Title: Affinity labelling of sorbitol dehydrogenase from sheep liver with alpha-bromo-

A:Reference number: S29379; MUID:93170323

A:Accession: S29379

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-18 <REI>

C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

-Query Match

19.1%; Score 30.5; DB 2; Length 18;

Best Local Similarity 38.5%; Pred. No. 6.1e+02;

Matches 5; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 16 GVEGSLH-WSYG 27

|:|:|:|:

Db 5 GIXGSDVHYWQH 17

RESULT 14

I39461

anti-angiotensin, hypothetical - human (fragment)

N:Alternate names: hypothetical angiotensin receptor antagonist, angiotensin mRNA comple

C:Species: Homo sapiens (man)

C:Date: 15-Feb-1996 #sequence_revision 15-Feb-1996 #text_change 20-Apr-2000

C:Accession: I39461

R:Moore, G.J.; Ganter, R.C.; Franklin, K.J.

Biochem. Biophys. Res. Commun. 160, 1387-1391, 1989

A:Title: Angiotensin 'antipeptides': (-)messenger RNA complementary to human angiotensin

A:Reference number: I39461; MUID:89273605

A:Accession: I39461

A:Molecule type: mRNA

A:Residues: 1-18 <MOO>

A:Cross-references: GB:M26228; NID:gl78641; PIDN:AAA35530.1; PID:gl78642

A:Note: this sequence is the conceptual translation of an nucleotide sequence complement

Query Match

18.8%; Score 30; DB 4; Length 18;

Best Local Similarity 62.5%; Pred. No. 7.2e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 EIKGVIVH 12

|:|:|:|

Db 6 EVEGVVH 13

RESULT 15

PS0188

superoxide dismutase (EC 1.15.1.1) (Cu-Zn), chloroplast - rice (fragment)

C:Species: Oryza sativa (rice)

C:Date: 05-Mar-1993 #sequence_revision 05-Mar-1993 #text_change 05-Mar-1999

C:Accession: PS0188

R:Kawakami, T.; Tsugita, A.

submitted to JIPID, June 1991

A:Reference number: PS0187

A:Accession: PS0188

A:Molecule type: protein

A:Residues: 1-20 <RAW>

A:Experimental source: leaf

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxye

C:Keywords: chloroplast; copper; metalloprotein; oxidoreductase; zinc

Query Match 18.8%; Score 30; DB 2; Length 20;

Best Local Similarity 41.2%; Pred. No. 8e+02;

Matches 7; Conservative 5; Mismatches 3; Indels 2; Gaps 1;

QY 1 KLLSEIKGVIVHRLGV 17

|:|:|:|

Db 4 KAVAILKG--THQVEGV 18

Search completed: October 10, 2002, 16:49:59

Job time : 12.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:41:22 ; Search time 9.5 Seconds
(without alignments)
126.348 Million cell updates/sec

Title: US-09-848-834A-9
Perfect score: 160
Sequence: 1 KLLSEIKGVVHRLEGVEGSLHWSYGLRXPX 31

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 1969
Minimum DB seq length: 0
Maximum DB seq length: 31

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	30.0	10	1 GON1_ALLMI	P37041 alligator m
2	41	25.6	10	1 GON1_CLUPA	P81749 clupea pall
3	39	24.4	10	1 GON3_ONCKE	P20367 oncorhynch
4	34	21.2	10	1 GON2_CHICK	P37043 gallus gall
5	34	21.2	10	1 GONL_SOUAC	P27429 squalus aca
6	32	20.0	26	1 PCW4_PACGO	P82426 pachycondyl
7	31	19.4	10	1 GON3_PETMA	P30948 petromyzon
8	31	19.4	24	1 CT31_LITCI	P81851 litoria cit
9	29.5	18.4	24	1 COXJ_SHEEP	Q9tr30 ovis aries
10	29	18.1	17	1 PC24_BRANA	P81097 brassica na
11	29	18.1	31	1 SODC_STRHE	P81163 striga herm
12	28.5	17.8	24	1 RS13_THETH	P80377 thermus aqu
13	28	17.5	28	1 PETL_CVAPA	P48102 cyanophora
14	27.5	17.2	24	1 AMAA_BACTR	P37356 bacillus th
15	27.5	17.2	29	1 COXK_SHEEP	Q9tr28 ovis aries
16	27	16.9	17	1 UP37_UPEMJ	P82044 uperoleia m
17	27	16.9	18	1 SODM_MYCHA	P80582 mycobacteri
18	27	16.9	20	1 ATP4_SPIOL	P80085 spinacia ol
19	26	16.2	10	1 GON1_CHEPR	P80677 chelyosoma
20	26	16.2	14	1 FIBA_HORSE	P14452 equus cabal
21	26	16.2	22	1 AOPA_MOUSE	Q64133 mus musculu
22	26	16.2	30	1 TL29_SPIOL	P81833 spinacia ol
23	25	15.6	8	1 ALI1_CYDPO	P82152 cydia pomon
24	25	15.6	10	1 GRP_RANRI	P32660 rana ridibu
25	25	15.6	17	1 TL09_SPIOL	P82671 spinacia ol
26	25	15.6	23	1 COXJ_ONCMY	P80333 oncorhynch
27	25	15.6	23	1 GRP_ONCMY	Q9ps30 oncorhynch
28	25	15.6	23	1 SODP_PICAB	P29427 picea abies
29	25	15.6	23	1 UHA4_HUMAN	P49289 homo sapien
30	25	15.6	25	1 CR14_LITGI	P56229 litoria gil
31	25	15.6	25	1 GRP_SCVCA	P09472 scyllorhinu
32	25	15.6	27	1 GRP_CANFA	P08989 canis famill
33	25	15.6	27	1 GRP_CHICK	P01295 gallus gall

RESULT 1				
ID	GONI_ALLMI	STANDARD;	PRT;	10 AA.
AC	P37041; P20407;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)			
DE	(Luliberin I).			
OS	Alligator mississippiensis (American alligator).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Crocodylidae; Alligatorinae; Alligator.			
OX	NCBI_TaxID=8496;			
RN	[1]			
RP	SEQUENCE.			
TX	TISSUE=Brain;			
RX	MEDLINE=91352338; PubMed=1882082;			
RA	Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,			
RA	Lance V., Swanson P., Rivier J.E., Sherwood N.M.;			
RT	"Primary structure of two forms of gonadotropin-releasing hormone			
RT	from brains of the American alligator (Alligator mississippiensis).";			
RL	Regul. Pept. 33:105-116(1991).			
CC	-!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- SIMILARITY: BELONGS TO THE GnRH FAMILY.			
DR	PIR; A60066; BHAQ1.			
DR	InterPro; IPR002012; GnRH.			
DR	Pfam; PF00446; GnRH; 1.			
DR	PROSITE; PS00473; GnRH; 1.			
KW	Hormone; Amidation; Hypothalamus.			
FT	MOD_RES 1			
FT	MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID.			
FT	MOD_RES 10 10 AMIDATION.			
SQ	SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;			
Query Match 30.0%; Score 48; DB 1; Length 10;				
Best Local Similarity 87.5%; Pred. No. 0.45;				
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;				
Qy	23 HWSYGLRP 30			
Db	2 HWSYGLQP 9			
RESULT 2				
GONI_CLUPA				
ID	GONI_CLUPA	STANDARD;	PRT;	10 AA.
AC	P81749;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH)			
DE	(Luliberin I).			
GN	GNRH1.			
OS	Clupea pallasii (Pacific herring).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;			

P01294 sus scrofa
P31886 alligator m
P00728 bacillus st
P81819 carcinus ma
P80575 streptomyc
P28524 hordeum vul
P54712 canis famill
P37155 felis silve
P81470 styela clav
P29428 picea abies
P33163 thermoleoph
P82421 pachycondyl

```

OC Clupeinae; Clupea.
OX NCBI_TaxID=30724;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carlsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
  Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
  hormones, including a novel form, from an ancient teleost, herring.";
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1105 MW; 284B20B72871F5A3 CRC64;

  Query Match      25.6%; Score 41; DB 1; Length 10;
  Best Local Similarity 75.0%; Pred. No. 4.5; Indels 1;
  Matches 6; Conservative 1; Mismatches 1; Gaps 0;

QY 23 HWSYGLRP 30
Db 2 HWSHGLSP 9
  |||||
  |||||

RESULT 3
GON3_ONCKE
ID GON3_ONCKE STANDARD; PRT; 10 AA.
AC P20367; P81751;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GnRH-III) (LH-
  RH III) (Luliberin III).
GN GnRH3.
OS Oncorhynchus keta (Chum salmon), and
  Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=O.keta;
RX MEDLINE=83195140; PubMed=6341999;
RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT "Characterization of a teleost gonadotropin-releasing hormone.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
RN [2]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carlsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
  Chang J.P., Rivier J.E., Sherwood N.M.;
RT "Primary structure and function of three gonadotropin-releasing
  hormones, including a novel form, from an ancient teleost, herring.";
RL Endocrinology 141:505-512(2000).
CC -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
  FOLLICLE-STIMULATING HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR PIR: A21114; A21114.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.

```

```

FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

  Query Match      24.4%; Score 39; DB 1; Length 10;
  Best Local Similarity 75.0%; Pred. No. 8.8;
  Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 23 HWSYGLRP 30
Db 2 HWSYGLWP 9
  |||||
  |||||

RESULT 4
GON2_CHICK
ID GON2_CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GnRH-II)
  (LH-RH II) (Luliberin II).
OS Gallus gallus (Chicken).
OS Alligator mississippiensis (American alligator),
  Squelus acanthias (Spiny dogfish),
OS Hydrolagus colliei (Spotted ratfish) (Pacific ratfish), and
  Clupea pallasii (Pacific herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN [1]
RP SEQUENCE.
RC SPECIES=Chicken; TISSUE=Hypothalamus;
RX MEDLINE=84222059; PubMed=6427779;
RA Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
  Matsuo H.;
RT "Identification of the second gonadotropin-releasing hormone in
  chicken hypothalamus: evidence that gonadotropin secretion is
  probably controlled by two distinct gonadotropin-releasing hormones
  in avian species.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=A.mississippiensis; TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
  Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone
  from brains of the American alligator (Alligator mississippiensis).";
RL Regul. Pept. 33:105-116(1991).
RN [3]
RP SEQUENCE.
RC SPECIES=S.acanthias; TISSUE=Brain;
RX MEDLINE=92335300; PubMed=1631133;
RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
  Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
  dogfish brain provides insight into GnRH evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN [4]
RP SEQUENCE.
RC SPECIES=H.colliei; TISSUE=Brain;
RX MEDLINE=91340067; PubMed=1678723;
RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
  Lee T.;
RT "Primary structure of gonadotropin-releasing hormone from the brain
  of a holoccephalan (ratfish: Hydrolagus colliei).";
RL Gen. Comp. Endocrinol. 82:152-161(1991).
RN [5]
RP SEQUENCE, AND FUNCTION.
RC SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX MEDLINE=20114351; PubMed=10650929;
RA Carlsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,

```

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
DR EMBL; AF003538; AAF49837.1; -.
DR FlyBase; FBgn0023095; caps.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR00372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 11.
DR Pfam; PF01463; LRRCT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 3.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_typ; 2.
SQ SEQUENCE 540 AA; 58911 MW; 203EBCD8FDC93266 CRC64;

Query Match 30.2%; Score 49; DB 5; Length 540;

Best Local Similarity 42.3%; Pred. No. 70;

Matches 11; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 6 GLRPGSGPSLQYIKANSKEFIGITEL 31

l:|:| |||||:|:|

Db 517 GVHPAAGG--YPYIAGNSRMIPVTEL 540

Search completed: October 10, 2002, 16:10:04

Job time : 15.0556 secs

```

RA Waterston R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: U08041; AAB37940.1; -.
DR InterPro: IPR000050; PID_domain.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR SMART: SM00462; PTB; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PSS0002; SH3; 1.
SQ SEQUENCE 315 AA; 36385 MW; B8572746211CFAAC CRC64;

Query Match 30.2%; Score 49; DB 5; Length 315;
Best Local Similarity 71.4%; Pred. No. 39;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 2 HWSYGV--LRPGSSG 13
||||| |||||
Db 79 HWSYGTNLTGTQSG 92

RESULT 13
Q50210
ID Q50210 PRELIMINARY; PRT; 481 AA.
AC Q50210;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE REVERSE TRANSCRIPTASE.
OS Melittangium lichenicola.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Cystobacteraceae; Melittangium.
OX NCBI_TaxID=45;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95095984; PubMed-7798147;
RA Rice S.A., Lamson B.C.;
RT "Phylogenetic comparison of retion elements among the myxobacteria:
RT evidence for vertical inheritance.";
RL J. Bacteriol. 177:37-45(1995).
DR EMBL: L36722; AAA66173.1; -.
DR InterPro: IPR000123; RNA_DNAPolys.
DR Pfam: PF00078; rvt; 1.
DR PRINTS; PR00866; RNADNAPOLMS.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 481 AA; 52786 MW; 3C4B51A562C82078 CRC64;

Query Match 30.2%; Score 49; DB 2; Length 481;
Best Local Similarity 39.1%; Pred. No. 62;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 8 RPGSSGSPSLQYIKANSKFIGITE 30
||| ||| ||| ||| |||
Db 437 KPGREGESLQKGMFAFIHMTD 459

RESULT 14
O96671
ID O96671 PRELIMINARY; PRT; 532 AA.
AC O96671;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE CAPRICIOUS.
GN CAPS OR CG11282.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CANTON-S;

```

```

RX MEDLINE-98306205; PubMed-9641918;
RA Shishido E., Takeichi M., Nose A.;
RT "Drosophila taysake formation: regulation by transmembrane protein
RT with Leu-rich repeats, CAPRICIOUS.";
RL Science 280:2118-2121(1998).
DR EMBL: AF102878; AAC78144.1; -.
DR FlyBase; FBgn0023095; caps.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_Out.
DR InterPro: IPR003591; LRR_tyr.
DR Pfam; PF00560; LRR; 11.
DR PRINTS; PR01463; LRRCT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00370; LRR; 3.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 2.
SQ SEQUENCE 532 AA; 57907 MW; 596AC5691C5D1723 CRC64;

Query Match 30.2%; Score 49; DB 5; Length 532;
Best Local Similarity 42.3%; Pred. No. 69;
Matches 11; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 6 GLRPGSSGSPSLQYIKANSKFIGITE 31
|:|:| ||| ||| ||| |||
Db 509 GVHPAAGG--YPIAGNSRMIPVTEL 532

RESULT 15
Q9VU53
ID Q9VU53 PRELIMINARY; PRT; 540 AA.
AC Q9VU53;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE CAPS PROTEIN.
GN CAPS OR CG11282.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

```

DT	01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DN	F26K24.25 PROTEIN (GENOMIC DNA, CHROMOSOME 3, PI CLONE: MEC18).
GE	F26K24.25.
GN	Arabidopsis thaliana (Mouse-ear cross).
OS	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX	NCBI_TaxID=3702;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. COLUMBIA;
RA	Liu X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA	Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA	Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT	"Arabidopsis thaliana chromosome III BAC F26K24 genomic sequence.";
RL	submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=COLUMBIA;
RA	Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=COLUMBIA;
RX	PubMed-10907853;
RA	Nakamura Y.;
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT	Sequence features of the regions of 4,251,695 bp covered by ninety pl
RT	TAC and BAC clones.";
RL	DNA Res. 7:217-221(2000).
DR	EMBL; AC016795; AAF23212.1; -.
DR	EMBL; AP002040; BAB03106.1; -.
SQ	SEQUENCE 1331 AA; 146497 MW; 2979FA0E8975143E CRC64;
Query Match	30.9%; Score 50; DB 10; Length 1331;
Best Local Similarity	30.4%; Pred.No.1.3e+02;
Matches	7; Conservative 9; Mismatches 7; Indels 0; Gaps
QY	4 SYGLRPGSGSPSLQYIKANSKFI 26
	I I I I I I I I I I I I I I I I :
Db	904 SYLKPGGETGSKMELVRVGVNEHV 926
RESULT 10	
Q90W09	
ID	Q90W09 PRELIMINARY; PRT; 54 AA.
QC	Q90W09;
DT	01-DEC-2001 (TEMBLrel. 19, Created)
DT	01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DE	01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE	GONADOTROPIN-RELEASING HORMONE PRECURSOR II ISOFORM D
DE	(GONADOTROPIN-RELEASING HORMONE PRECURSOR II ISOFORM C).
OC	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri);
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC	Percanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX	NCBI_TaxID=8022;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=TESTIS;
RX	MEDLINE=21232987; PubMed-11335940;
RA	Uzbekova S., Ferriere F., Guiguen Y., Bailhache T., Breton B.,
RA	Lareyre J.;
RT	"Stage-dependent and alternative splicing of sGrRH messengers in
RT	rainbow trout testis during spermatogenesis.";
RL	Mol. Reprod. Dev. 59:1-10(2001).
DR	EMBL; AF269107; AAK54679.1; -.
DR	EMBL; AF269106; AAK54678.1; -.
SQ	SEQUENCE 54 AA; 5963 MW; 06BF365F658E08B7 CRC64;
Query Match	30.6%; Score 49.5; DB 13; Length 54;
Best Local Similarity	42.9%; Pred. No. 4.9;
Matches	12; Conservative 4; Mismatches 9; Indels 3; Gaps

```
FT NON_TER 1 1
FT NON_TER 87 87
SQ SEQUENCE 87 AA; 9871 MW; 0D2463533D96782A CRC64;

Query Match 32.1%; Score 52; DB 13; Length 87;
Best Local Similarity 88.9%; Pred. No. 3.5;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 22 HWSYGLSPG 30

RESULT 6
Q9RDQ1
ID Q9RDQ1 PRELIMINARY; PRT; 1545 AA.
AC Q9RDQ1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PUTATIVE SECRETED PROTEIN.
GN SC4A7.11.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Kinashi H., Hopwood D.A.;
RL "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RT Mol. Microbiol. 21:77-96(1996).
DR EMBL; ALI33423; CAB62715.1; -.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00652; Ricin_B_lectin; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS50231; RICIN_B_LLECTIN; 1.
SQ SEQUENCE 1545 AA; 161520 MW; 81EF325143593ABA CRC64;

Query Match 31.5%; Score 51; DB 2; Length 1545;
Best Local Similarity 31.9%; Pred. No. 1.1e+02;
Matches 15; Conservative 4; Mismatches 10; Indels 18; Gaps 2;

QY 2 HWSYGLR-----PGSSGPSL-QYIKANSKFTIGITE 30
DB 1408 HWARYRTESEYFEGDHNVAVKVPGWNPRTGDFLIANSKFSHSE 1454

RESULT 7
O88902
ID O88902 PRELIMINARY; PRT; 1494 AA.
AC O88902;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE TD14.
GN PTP-TD14.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

```
OX NCBI_TaxID=10116;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98361981; PubMed=9694860;
RA Cao L., Zhang L., Ruiz-Lozano P., Yang Q., Chien K.R., Graham R.M.,
RA Zhou M.;
RT "A novel putative protein-tyrosine phosphatase contains a BRO1-like domain and suppresses Ha-ras-mediated transformation.";
RL J. Biol. Chem. 273:21077-21083(1998).
DR EMBL; AF077000; AAC62959.1; -.
DR HSSP; Q06124; 2SHP.
DR InterPro; IPR000367; TYR_phosphatase.
DR InterPro; IPR000242; Tyr_prot_phptase.
DR Pfam; PF00102; Y_phosphatase; 1.
DR SMART; SM00194; PTPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolase.
SQ SEQUENCE 1494 AA; 162931 MW; 8F42DF6CD40D0E90 CRC64;

Query Match 31.2%; Score 50.5; DB 11; Length 1494;
Best Local Similarity 36.4%; Pred. No. 1.3e+02;
Matches 12; Conservative 6; Mismatches 12; Indels 3; Gaps 1;

QY 2 HWSYGLRPGSSGPSLQYIKR---NSKFIGITEL 31
DB 593 HFSPPGPGSTGTPATHYLSGPLPPGYSGPTQL 625

RESULT 8
Q28848
ID Q28848 PRELIMINARY; PRT; 224 AA.
AC Q28848;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE TESTIS-DETERMINING PROTEIN (FRAGMENT).
OS Sminthopsis macroura (Stripe-faced Dunnart).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Dasyuromorphia; Dasyuridae; Sminthopsis.
OX NCBI_TaxID=9302;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=94173941; PubMed=8127908;
RA Foster J.W., Graves J.A.;
RT "An SRV-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:1927-1931(1994).
DR EMBL; S69429; AAB30154.2; -.
DR HSSP; Q05066; 1HRY.
DR InterPro; IPR000910; HMG_12_box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
FT NON_TER 224 224
SQ SEQUENCE 224 AA; 24128 MW; 96F284710D3B8A91 CRC64;

Query Match 30.9%; Score 50; DB 6; Length 224;
Best Local Similarity 64.3%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 SYGLRPGSSGPSLQ 17
DB 185 SYGQHPGMNGPOLQ 198

RESULT 9
Q9SF03
ID Q9SF03 PRELIMINARY; PRT; 1331 AA.
AC Q9SF03;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
```



```

GN NRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21102951; PubMed=11170016;
RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GnRH1 and GnRH2
RT precursors from bullfrog (Rana catesbeiana).";
RL J. Exp. Zool. 289:190-201(2001).
DR EMBL; AF188754; AAL05972.1; -.
SQ SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;

Query Match 36.7%; Score 59.5; DB 13; Length 90;
Best Local Similarity 47.1%; Pred. No. 0.28; Mismatches 9; Indels 5; Gaps 3;
Matches 16; Conservative 4;

QY 2 HWSYGLRPGSG--PSLQ--YIKANSKFIGITEL 31
Db 26 HWSYGLRPGKREVSLESQAEVNPNE-VSFTTEL 58
|||||
| | | | |

RESULT 3
ID Q9PRH0 PRELIMINARY; PRT; 91 AA.
AC Q9PRH0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE GONADOTROPIN PRECURSOR (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-
DE RH) (LULIBERIN).
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguilloidei;
OC Anguillidae; Anguilla.
OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Okubo K., Suetake H., Aida K.;
RT "Expression of two gonadotropin-releasing hormone (GNRH) precursor
RT genes in various tissues of the Japanese eel and evolution of GNRH.";
RL Zool. Sci. 16:471-478(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Okubo K., Suetake H., Aida K.;
RT "A splicing variant for the prepro-mammalian gonadotropin-releasing
RT hormone (prepro-mGNRH) mRNA is present in the brain and various
RT peripheral tissues of the Japanese eel.";
RL Zool. Sci. 16:645-651(1999).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AB026989; BAA82608.1; -.
DR EMBL; AB026991; BAA83597.1; -.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBRN1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone; Signal.
FT SIGNAL 1 22
FT CHAIN 23 32 MGNRH.
FT CHAIN 33 91 GNRH ASSOCIATED PEPTIDE.
SQ SEQUENCE 91 AA; 9893 MW; BA15C9DC08434A7B CRC64;

Query Match 36.4%; Score 59; DB 13; Length 91;
Best Local Similarity 66.7%; Pred. No. 0.33; Mismatches 4; Indels 2; Gaps 1;
Matches 12; Conservative 0;

QY 2 HWSYGLRPGSG--SGPSLQ 17
Db 24 HWSYGLRPGGKRGADSLQ 41
|||||
| | | | |

RESULT 4
Q92YR6 PRELIMINARY; PRT; 374 AA.
ID Q92YR6;
AC Q92YR6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE MUCONATE CYCLOISOMERASE (EC 5.5.1.1).
GN SMAIL461.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL; AE007266; AAK65455.1; -.
KW Isomerase; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 374 AA; 40999 MW; 0369AA67790B23D6 CRC64;

Query Match 33.0%; Score 53.5; DB 16; Length 374;
Best Local Similarity 44.8%; Pred. No. 10; Mismatches 10; Indels 1; Gaps 1;
Matches 13; Conservative 5;

QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITE 30
Db 21 HWSYGLRPGSFAVNLIEADDTGTVGIGE 48
|||||
| | | | |

RESULT 5
Q9YI26 PRELIMINARY; PRT; 87 AA.
ID Q9YI26;
AC Q9YI26;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE (LULIBERIN) (FRAGMENT).
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Nabissi M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR EMBL; AF046801; AAD02427.1; -.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Amidation; Hormone.

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 12.0556 Seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSGLRPGSGPSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	76	46.9	1310	2 Q93N27	Q93n27 clostridium
2	59.5	36.7	90	13 Q90Y63	Q90y63 rana catesb
3	59	36.4	91	13 Q9PRH0	Q9prh0 anguilla ja
4	53.5	33.0	374	16 Q92Y86	Q92yr6 rhizobium m
5	52	32.1	87	13 Q9YI26	Q9yi26 sparus aura
6	51	31.5	1545	2 Q9RDQ1	Q9rdq1 streptomyce
7	50.5	31.2	1494	11 Q88902	Q88902 rattus norv
8	50	30.9	224	6 Q28848	Q28848 sminthopsis
9	50	30.9	1331	10 Q9SF03	Q9sf03 arabidopsis
10	49.5	30.6	54	13 Q90W09	Q90w09 oncorhynch
11	49.5	30.6	256	16 P72950	P72950 synechocyst
12	49	30.2	315	5 P91045	P91045 caenorhabdi
13	49	30.2	481	2 Q50210	Q50210 melittangiu
14	49	30.2	532	5 Q96671	Q96671 drosophila
15	49	30.2	540	5 Q9VU53	Q9vu53 drosophila
16	49	30.2	1047	3 Q9P6E3	Q9p6e3 neurospora

17	48	29.6	91	13 Q9DGC8	Q9dgc8 oryzias lat
18	48	29.6	104	11 Q9JL82	Q9j182 mus musculu
19	48	29.6	145	6 Q28568	Q28568 ovis aries
20	48	29.6	146	17 Q27851	Q27851 methanother
21	48	29.6	306	10 Q9SDN3	Q9sdn3 prunus dulc
22	48	29.6	361	5 Q94176	Q94176 caenorhabdi
23	48	29.6	518	5 Q17807	Q17807 caenorhabdi
24	48	29.6	4824	5 Q95YM1	Q95ym1 procambaru
25	48	29.6	4928	2 Q9ALM3	Q9alm3 saccharopol
26	48	29.6	17352	5 Q95YM2	Q95ym2 procambaru
27	47.5	29.3	88	13 Q9PSY9	Q9psy9 sparus aura
28	47.5	29.3	90	13 Q9DDA9	Q9dda9 oryzias lat
29	47.5	29.3	201	2 Q9ZFA0	Q9zfa0 streptomyce
30	47.5	29.3	264	16 Q31775	Q31775 bacillus su
31	47.5	29.3	505	16 Q92Y56	Q92y56 rhizobium m
32	47.5	29.3	515	5 Q95SL0	Q95sl0 drosophila
33	47.5	29.3	604	2 Q9RIY1	Q9riy1 streptomyce
34	47.5	29.3	794	2 P72249	P72249 rhodobacter
35	47.5	29.3	1037	5 Q9W2D5	Q9w2d5 drosophila
36	47.5	29.3	1067	15 P89904	P89904 chimpanzee
37	47	29.0	444	16 P71718	P71718 mycobacteri
38	47	29.0	634	10 Q94LW6	Q94lw6 arabidopsis
39	47	29.0	1146	13 Q90584	Q90584 gallus gall
40	47	29.0	1970	5 Q9VQU8	Q9vqu8 drosophila
41	46.5	28.7	560	15 Q90S55	Q90s55 human immun
42	46	28.4	220	9 Q37946	Q37946 bacterioph
43	46	28.4	246	5 Q9XVQ5	Q9xvq5 caenorhabdi
44	46	28.4	252	9 Q9XJE8	Q9xje8 lactococcus
45	46	28.4	356	6 Q9BGQ0	Q9bgq0 macaca fasc

ALIGNMENTS

RESULT 1

Q93N27 Q93N27 PRELIMINARY; PRT; 1310 AA.
AC Q93N27;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TETANUS TOXIN (FRAGMENT).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA Shumin Z., Dianliang L.;
RT "Cloning and sequence analysis of tetanus toxin gene."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389424; AAK72964.2; -.
FT NON_TER 1 1
FT NON_TER 1310 1310
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 46.9%; Score 76; DB 2; Length 1310;
Best Local Similarity 93.8%; Pred. No. 0.017;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LQYIKANSKFIGITEL 31
:|||||
DB 830 MQYIKANSKFIGITEL 845

RESULT 2

Q90Y63 Q90Y63 PRELIMINARY; PRT; 90 AA.
AC Q90Y63;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GONADOTROPIN-RELEASING HORMONE.


```

RP SEQUENCE FROM N.A.
RA CHOW M.M., Kight K.E., Gothliff Y., Alok D., Zohar Y.;
RT "Multiple GnRHs present in a teleost species are encoded by separate
RT genes: analysis of the sbGnRH and cGnRH-II genes from the striped
RT bass, Morone saxatilis.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF056314; AAD03817.1; -.
DR InterPro: IPR002012; GnRH.
DR Pfam: PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 95 PROGONADOLIBERIN I.
FT PEPTIDE 23 32 GONADOLIBERIN I.
FT PEPTIDE 36 95 GnRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP)
FT (BY SIMILARITY).
SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match 32.1%; Score 52; DB_1; Length 95;
Best Local Similarity 88.9%; Pred. NO. 0.64;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
   |||||
Db 24 HWSYGLSPG 32

```

Search completed: October 10, 2002, 16:06:34
Job time : 5.0406 secs

```

RP SYNTHESIS OF 24-33.
RX MEDLINE=82265777; PubMed=7050118;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RT hormone. I. Structural determination on partially purified
RT material.";
RT J. Biol. Chem. 257:10722-10728(1982).
RC CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X69491; CAA49246.1; -.
CC PIR; S33507; S33507.
CC InterPro; IPR002012; GnRH.
CC Pfam; PF00446; GnRH; 1.
CC PRINTS; PR01541; GONADOLIBERN.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal.
CC CHAIN 1 23 PROGONADOLIBERIN I.
CC PEPTIDE 24 32 GONADOLIBERIN I.
CC PEPTIDE 24 33 GONADOLIBERIN I.
CC PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.
CC MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
CC MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
CC SEQUENCE 92 AA; 10206 MW; 61AEB7EBAF508B6A CRC64;

Query Match 33.3%; Score 54; DB 1; Length 92;
Best Local Similarity 88.9%; Pred. No. 0.31;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||:|
DB 25 HWSYGLQPG 33

RESULT 14
GONL HAPBU STANDARD; PRT; 94 AA.
AC P51918; O93387;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
DE (LH-RH I) (Luliberin I).
GN GNRH1.
OS Haplochromis burtoni (Burton's mouthbrooder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Astotitilapia.
OX NCBI_TaxID=8153;

SEQUENCE FROM N.A.
RA MEDLINE=95396797; PubMed=7667296;
RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;
RT "Three gonadotropin-releasing hormone genes in one organism suggest
RT novel roles for an ancient peptide.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
RN [2]
SEQUENCE FROM N.A.
RA MEDLINE=99061842; PubMed=9843638;
RA White R.B., Fernald R.D.;
RT "Ontogeny of gonadotropin-releasing hormone (GnRH) gene expression

```

```

RT reveals a distinct origin for GnRH-containing neurons in the
RT midbrain.";
RL Gen. Comp. Endocrinol. 112:322-329(1998).
RN [3]
RP SEQUENCE OF 23-32.
RC TISSUE-Pituitary;
RX MEDLINE=95372591; PubMed=7644702;
RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
RA White S.A., Francis R.C., Fernald R.D., Licht P., Warby C.,
RA Sherwood N.M.;
RT "Primary structure of solitary form of gonadotropin-releasing hormone
RT (GnRH) in cichlid pituitary; three forms of GnRH in brain of cichlid
RT and pumpkinseed fish.";
RL Regul. Pept. 57:43-53(1995).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-
CC GONADAL AXIS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOPHYSAL AXONS.
CC -!- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U31865; AAC59691.1; -.
CC EMBL; AF076961; AAC27716.1; -.
CC InterPro; IPR002012; GnRH.
CC Pfam; PF00446; GnRH; 1.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal; Multigene family.
CC SIGNAL 1 22
CC CHAIN 23 94 PROGONADOLIBERIN I.
CC PEPTIDE 23 32 GONADOLIBERIN I.
CC PEPTIDE 36 94 GnRH-ASSOCIATED PEPTIDE I (POTENTIAL).
CC MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.
CC MOD_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP).
CC CONFLICT 86 94 ENGRFTKK -> KMDTGHSRNERFL (IN REF. 1).
CC SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;

Query Match 32.1%; Score 52; DB 1; Length 94;
Best Local Similarity 88.9%; Pred. No. 0.63;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||:|
DB 24 HWSYGLSPG 32

RESULT 15
GONL MORSA STANDARD; PRT; 95 AA.
AC O73812;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
DE (LH-RH I) (Luliberin I).
GN GNRH1.
OS Morone saxatilis (Striped bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Moronidae; Morone.
OX NCBI_TaxID=34816;
RN [1]

```

DE	(Luliberin I).
OS	Alligator mississippiensis (American alligator).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX	NCBI_TaxID=8496;
RN	[1]
RP	SEQUENCE.
RC	TISSUE=Brain;
RX	MEDLINE=91352338; PubMed=1882082;
RA	Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA	Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT	"Primary structure of two forms of gonadotropin-releasing hormone
RT	from brains of the American alligator (Alligator mississippiensis).
RL	Regul. Pept. 33:105-116(1991).
CC	- - FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC	- - SUBCELLULAR LOCATION: Secreted.
CC	- - SIMILARITY: BELONGS TO THE GnRH FAMILY.
DR	PIR; A60066; RHAQ1.
DR	InterPro; IPR002012; GnRH.
DR	Pfam; PF00446; GnRH; 1.
DR	PROSITE; PS00473; GnRH; 1.
KW	Hormone; Amidation; Hypothalamus.
FT	MOD_RES 1 PYRROLIDONE CARBOXYLIC ACID.
FT	MOD_RES 10 10 AMIDATION.
SQ	SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;
Query Match 33.3%; Score 54; DB 1; Length 10;	
Best Local Similarity 88.9%; Pred.No. 0.025; Indels 0; Gaps	
Matches 8; Conservative 1; Mismatches 0;	
QY	2 HWSYGLRPG 10
	:
DB	2 HWSYGLQPG 10
RESULT 13	
CON1_CHICK	
ID	GON1_CHICK STANDARD; PRT; 92 AA.
AC	P37042; P20407;
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-JUN-1994 (Rel. 29, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE	(Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE	hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide 1].
OC	Gallus gallus (Chicken).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae
OX	NCBI_TaxID=9031;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=WHITE LEGHORN;
RX	MEDLINE=94059355; PubMed=7902095;
RA	Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
RT	"Characterization of the chicken pre-gonadotropin-releasing
RT	hormone-I gene";
RN	J. Mol. Endocrinol. 11:19-29(1993).
RN	[2]
RP	SEQUENCE OF 24-33.
RC	TISSUE=Hypothalamus;
RX	MEDLINE=82265778; PubMed=7050119;
RA	King J.A., Millar R.P.;
RT	"Structure of chicken hypothalamic luteinizing hormone-releasing
RT	hormone. II. Isolation and characterization.;"
RN	J. Biol. Chem. 257:10729-10732(1982).
RN	[3]
RP	SEQUENCE OF 24-33.
RC	TISSUE=Hypothalamus;
RA	King J.A., Millar R.P.;
RT	"Structure of avian hypothalamic gonadotropin-releasing hormone.;"
RN	S. Afr. J. Sci. 78:124-125(1982).
RN	[4]

```

QY      2 HWSYGLRPG 10
Db      25 HWSYGLRPG 33

RESULT 9
GONI_RAT
ID GONI_RAT      STANDARD;      PRT;      92 AA.
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
DE I].
GN GnRH1 OR GnRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus.";
RL Science 235:1514-1517(1987).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS: IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -
DR EMBL; M12579; AAA41263.1; -
DR EMBL; M31670; AAA41264.1; -
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -

DR EMBL; M15528; -, NOT_ANNOTATED_CDS.
DR PIR; B26173; RHRTG.
DR PIR; A48410; A48410.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL      1      23      PROGONADOLIBERIN I.
FT CHAIN       24      92      GONADOLIBERIN I.
FT PEPTIDE     24      33      PROLACTIN RELEASE-INHIBITING FACTOR I.
FT PEPTIDE     37      92      APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE    26      26      ACTIVITY.
FT MOD_RES     24      24      PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES     33      33      AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE    92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match      35.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.078;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 HWSYGLRPG 10
Db      25 HWSYGLRPG 33

RESULT 10
GONI_TUPGB
ID GONI_TUPGB      STANDARD;      PRT;      92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
GN GnRH1 OR GnRH.
OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaia.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGnRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GnRH gene expression in a
RT placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U63326; AAB16837.1; -
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRNI.

```

RL Biochem./ Biophys. Res. Commun. 44:459-463(1971).
 RN [3]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72065376; PubMed=4942726;
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 phase method";
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
 RN [4]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72117544; PubMed=4946275;
 RA Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L32864; AAA31066.1; -;
 DR PIR; A01411; RHPGG.
 DR InterPro; IPR002012; GNRH.
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBRNI.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 91 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
 Query Match 35.8%; Score 58; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 I|||||
 Db 25 HWSYGLRPG 33
 RESULT 8
 GONL_HUMAN
 ID GONL_HUMAN STANDARD; PRT; 92 AA.
 AC P01148;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Progadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated
 DE peptide I].
 GN GNRH1 OR GNRH OR LHRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=89366682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 hormone gene";
 RL Nucleic Acids Res. 17:6403-6403(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 of gonadotropin-releasing hormone and prolactin release-inhibiting
 factor in human and rat.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85012739; PubMed=6090951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 releasing hormone";
 RL Nature 311:666-668(1984).
 RN [4]
 RP SEQUENCE OF 24-33;
 RX MEDLINE=83126573; PubMed=6760865;
 RA Tan L., Rousseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide
 biosynthesized in the human placenta.";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 CC Lutrepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm
 CC (Serono).
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X01059; CAA25526.1; -;
 DR EMBL; M12578; AAA35916.1; -;
 DR EMBL; X15215; CAA33285.1; -;
 DR PIR; A01410; RHHUG.
 DR PIR; A26173; A26173.
 DR PIR; S05308; S05308.
 DR MIN; 152760; -;
 DR InterPro; IPR002012; GNRH.
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBRNI.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pharmaceutical; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 92 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 FT CONFLICT 16 16 W -> S (IN REF. 3).
 SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;
 Query Match 35.8%; Score 58; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Conadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN=WESTERN RANGE; TISSUE=Hypothalamus;
RA Rodriguez R.E., Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-10.
RX MEDLINE=72094314; PubMed=4550508;
RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vale W., Guillemin R.;
RT "Primary structure of the ovine hypothalamic luteinizing hormone-
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-Edman degradation).";
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U02517; AAA03433.1; -.
DR PIR; A93780; RHSHG.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >61 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6828 MW; 63962AJAE319B9F0 CRC64;
Query Match 37.0%; Score 60; DB 1; Length 61;
Best Local Similarity 52.2%; Pred. No. 0.024;
Matches 12; Conservative 1; Mismatches 2; Indels 8; Gaps 1;
Qy 2 HWSYGLRPGSGPSLQYTKANSK 24
Db 2 HWSYGLRPGG-----KRNK 16
RESULT 4
GONI_MACMU STANDARD; PRT; 67 AA.
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Conadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S75918; AAB33096.1; -.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 5 BY SIMILARITY.
FT CHAIN 6 >67 PROGNADOLIBERIN I.
FT PEPTIDE 6 15 GONADOLIBERIN I.
FT PEPTIDE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
FT NON_TER 67 67 SIMILARITY).
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
Query Match 35.8%; Score 58; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 HWSYGLRPG 10
Db 7 HWSYGLRPG 15
RESULT 5
GONI_XENLA STANDARD; PRT; 89 AA.
AC P45656;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (Gnrh-I)
DE (LH-RH) (Luliberin I).
OS Xenopus laevis (African clawed frog).

EMBO J. 11:3577-3583(1992).
[7]
IDENTIFICATION OF SUBSTRATE.
Schlavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
Dasgupta B.R., Montecucco C.;
"Tetanus and botulinum-B neurotoxins block neurotransmitter release
by proteolytic cleavage of synaptobrevin.";
Nature 359:832-835(1992).
[8]
X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
MEDLINE=97475217; PubMed=9334741;
Umland T.C., Wingert L.M., Swaminathan S., Purey W.F., Schmidt J.J.,
Sax M.;
"Structure of the receptor binding fragment HC of tetanus
neurotoxin.";
Nat. Struct. Biol. 4:788-792(1997).
-1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77
BOND OF SYNAPTOBREVIN-2.
-1- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-1-PHE-77 BOND IN
SYNAPTOBREVIN.
-1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
AND ARE NON-TOXIC AFTER SEPARATION.
-1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
GANGLIOSIDE RECEPTORS.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X04436; AAA28033.1; -
EMBL; M12739; AAA23282.1; -
EMBL; X06214; CAA29584.1; -
PIR; A25689; BTCLTN.
PDB; 1AF9; 29-APR-98.
PDB; 1A8D; 14-OCT-98.
MEROPS; M27_001; -
InterPro; IPR000395; Bontoxilysin.
InterPro; IPR000130; Zn_Mtpeptidse.
Pfam; PF01742; Peptidase_M27; 1.
PRINTS; PR00760; BONTOXILYSIN.
ProDom; PD001963; Bontoxilysin; 1.
PROSITE; PS00142; ZINC_PROTEASE; 1.
Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
3D-structure.
INIT_MET 0
CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
ACT_SITE 233 233 BY SIMILARITY.
METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
TRANSMEM 226 246 POTENTIAL.
TRANSMEM 669 689 POTENTIAL.
DISULFID 438 466 INTERCHAIN.
DISULFID 1076 1092
SEQUENCE 1314 AA; 150550 MW; 134C3657133EF81D CRC64;
Query Match 46.9%; Score 76; DB 1; Length 1314;
Best Local Similarity 93.8%; Pred. No. 0.0031;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 16 LQYIKANSKFIGITEL 31

DB 828 MOYIKANSKFIGITEL 843
RESULT 2
GONI_MESAU
ID GONI_MESAU STANDARD; PRT; 63 AA.
AC 009163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; U91938; AAB51302.1; -
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >63 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY
SIMILARITY).
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT MOD_RES 1 1 ACTIVITY (BY SIMILARITY).
FT MOD_RES 10 10 PYRROLIDONE CARBOXYLIC ACID (BY
SIMILARITY).
FT MOD_RES 10 10 AMIDATION (6-11 PROVIDE AMIDE GROUP) (BY
SIMILARITY).
FT NON_TER 63 63
FT SEQUENCE 63 AA; 7370 MW; FC9499567677180 CRC64;
Query Match 37.7%; Score 61; DB 1; Length 63;
Best Local Similarity 46.2%; Pred. No. 0.018;
Matches 12; Conservative 3; Mismatches 3; Indels 8; Gaps 1;
QY 2 HWSYGLRPGSGPSLQYIKANSKFIG 27
IIIIIIIIII I:::
DB 2 HWSYGLRPGG-----KRAERLG 19
RESULT 3
GONI_SHEEP
ID GONI_SHEEP STANDARD; PRT; 61 AA.
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 4.0406 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSYGLRPGSSGSPSQYIKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	46.9	1314	1 TETX_CLOTE	P04958 clostridium
2	61	37.7	63	1 GON1_MESAU	O09163 mesocricetu
3	60	37.0	61	1 GON1_SHEEP	O28588 ovis aries
4	58	35.8	67	1 GON1_MACMU	P53247 macaca mula
5	58	35.8	89	1 GON1_XENLA	P45656 xenopus lae
6	58	35.8	90	1 GON1_MOUSE	P13562 mus muscula
7	58	35.8	91	1 GON1_PIG	P49921 sus scrofa
8	58	35.8	92	1 GON1_HUMAN	P01148 homo sapien
9	58	35.8	92	1 GON1_RAT	P07490 rattus norv
10	58	35.8	92	1 GON1_TUPGB	P05335 tupaia gliis
11	55.5	34.3	66	1 V884_BPML5	Q05301 mycobacteri
12	54	33.3	10	1 GON1_ALLMI	P37041 alligator m
13	54	33.3	92	1 GON1_CHICK	P37042 gallus gall
14	52	32.1	94	1 GON1_HAPBU	P51918 haplochromi
15	52	32.1	95	1 GON1_MORSA	O73812 morone saxa
16	52	32.1	95	1 GON1_PAGMA	P70074 pegrus majo
17	52	32.1	95	1 GON1_SPAAU	P51919 sparus aura
18	52	32.1	99	1 GON1_DICLA	O91a10 dicentrarch
19	52	32.1	575	1 ACEA_LYCES	P49297 lycopersico
20	50.5	31.2	90	1 GON3_DICLA	O91a09 dicentrarch
21	50	30.9	80	1 GON3_CLAGA	P33439 ciarias gar
22	50	30.9	90	1 GON8_RANDY	O91a02 rana dybows
23	50	30.9	249	1 PRA_MYCLE	P41484 mycobacteri
24	49	30.2	92	1 GON1_CAVPO	O54713 cavia porce
25	49	30.2	293	1 KHSE_PYRHO	O58814 pyrococcus
26	49	30.2	294	1 KHSE_PYRAB	O94297 pyrococcus
27	49	30.2	408	1 SEPR_THESR	P80146 thermus sp.
28	49	30.2	444	1 GARP_ECOLI	P42613 escherichia
29	49	30.2	485	1 RT16_MYXXA	P23072 myxococcus
30	48	29.6	390	1 YB12_SCHPO	P87167 schizosacch
31	47.5	29.3	89	1 GON3_PORNO	P51922 porichthys
32	47.5	29.3	90	1 GON3_HAPBU	P45652 haplochromi
33	47.5	29.3	90	1 GON3_PAGMA	P51921 pegrus majo

RESULT 1

ID	TETX_CLOTE	STANDARD;	PRT;	1314 AA.
AC	P04958;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).			
OS	Clostridium tetani.			
OG	Plasmid.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1513;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87053814; PubMed=3536478;			
RA	Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,			
RA	Weller U., Hudel M., Habermann E., Niemann H.;			
RT	"Tetanus toxin: primary structure, expression in E. coli, and			
RT	homology with botulinum toxins.";			
RL	EMBO J. 5:2495-2502(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CN3911;			
RX	MEDLINE=87040747; PubMed=3774547;			
RA	Fairweather N.F., Lyness V.A.;			
RT	"The complete nucleotide sequence of tetanus toxin.";			
RL	Nucleic Acids Res. 14:7809-7812(1986).			
RN	[3]			
RP	SEQUENCE OF 742-1314 FROM N.A.			
RX	MEDLINE=86085672; PubMed=3510187;			
RA	Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;			
RT	"Cloning, nucleotide sequencing, and expression of tetanus toxin			
RT	fragment C in Escherichia coli.";			
RL	J. Bacteriol. 165:21-27(1986).			
RN	[4]			
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.			
RX	MEDLINE=90201034; PubMed=2108021;			
RA	Kriegstein K., Henschen A., Weller U., Habermann E.;			
RT	"Arrangement of disulfide bridges and positions of sulfhydryl groups			
RT	in tetanus toxin.";			
RL	Eur. J. Biochem. 188:39-45(1990).			
RN	[5]			
RP	PARTIAL SEQUENCE.			
RX	MEDLINE=92037649; PubMed=1935979;			
RA	Kriegstein K.G., Henschen A.H., Weller U., Habermann E.;			
RT	"Limited proteolysis of tetanus toxin. Relation to activity and			
RT	identification of cleavage sites.";			
RL	Eur. J. Biochem. 202:41-51(1991).			
RN	[6]			
RP	IDENTIFICATION AS ZINC-PROTEASE.			
RX	MEDLINE=93010948; PubMed=1396558;			
RA	Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,			
RA	Montecucco C.;			
RT	"Tetanus toxin is a zinc protein and its inhibition of			
RT	neurotransmitter release and protease activity depend on zinc.";			

ALIGNMENTS

Db 2 HWSYGLQPG 10

RESULT 13

I50644

gonadoliberin I precursor - chicken

N:Alternate names: gonadotropin-releasing hormone I

C:Species: gallus gallus (chicken)

C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999

C:Accession: I50644; S33507

R:Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.

J. Mol. Endocrinol. 11, 19-29, 1993

A:Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene.

A:Reference number: I50644; MUID:94059355

A:Accession: I50644

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-92 <DU2>

A:Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA49246.1; PID:g311612

C:Genetics:

A:Introns: 47/3; 79/3

C:Superfamily: gonadoliberin

Query Match

Best Local Similarity 33.3%; Score 54; DB 2; Length 92;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

|||||:||

Db 25 HWSYGLQPG 33

RESULT 14

E95361

probable muconate cycloisomerase (EC 5.5.1.1) [imported] - Sinorhizobium meliloti (strain

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: E95361

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kallan, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: E95361

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK65455.1; PID:gl4523923; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kallan, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Sma1461

A:Genome: plasmid

C:Keywords: intramolecular lyase; isomerase

Query Match

Best Local Similarity 33.0%; Score 53.5; DB 2; Length 374;

Matches 13; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 2 HWSYGLRPGSGPSLOYIKANSKFGITE 30

|||||:| | | | | | | | | |

Db 21 HWSYGIRE-SFAVNLIEIEADDTGVIGE 48

RESULT 15

I50739

gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)

C:Species: Haplochromis burtoni

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: I50739

R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.

Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995

A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel rol

A:Reference number: I50739; MUID:95396797

A:Accession: I50739

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-98 <WHI>

A:Cross-references: EMBL:U31865; NID:g905398; PIDN:AAC59691.1; PID:g905399

C:Superfamily: gonadoliberin

Query Match

Best Local Similarity 32.1%; Score 52; DB 2; Length 98;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

|||||:||

Db 24 HWSYGLSPG 32

Search completed: October 10, 2002, 16:12:11

Job time : 8.29701 secs

A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA1263.1; PID:g204446
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone
A:Reference number: A48410; MUID:93105480
A:Accession: A48410
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <MAI>
A:Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
A:Experimental source: thymus
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)
C:Genetics:
A:Introns: 47/3; 79/3
C:Function:
A:Description: stimulates pituitary secretion of lutropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid; F:1-23/domain: signal sequence #status predicted <SIG>
F:1-23/Product: progonaoliberin #status predicted <PGN>
F:24-92/Product: gonadoliberin #status predicted <GLN>
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 35.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
|||||

RESULT 9
T25210
hypothetical protein B2J23.60 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
A:Accession: T52510
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, September 2000
A:Reference number: Z26053
A:Accession: T52510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <SCH>
A:Cross-references: EMBL:AL442164; GSPDB:GN00116; NCSP:B2J23.60
A:Experimental source: BAC clone B2J23; strain OR7A
C:Genetics:
A:Gene: NCSP:B2J23.60
A:Map position: 6
A:Introns: 349/1; 601/1

Query Match 34.6%; Score 56; DB 2; Length 719;
Best Local Similarity 45.0%; Pred. No. 6.3;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 WSYGLRPGSGPSLOYIKAN 22
DB 557 WSYGRPGSAGGLMSFVSAS 576
|||||

RESULT 10
S31029
gene 84 protein - Mycobacterium phage L5
C:Species: Mycobacterium phage L5
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
A:Accession: S31029
R:Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
Mol. Microbiol. 7, 407-417, 1993

A:Title: Superinfection immunity of mycobacteriophage L5: applications for genetic tr
A:Reference number: S30949; MUID:93211283
A:Accession: S31029
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-66 <DON>
A:Cross-references: EMBL:Z18946; NID:g15959; PIDN:CAA79460.1; PID:e59702; PID:g579152
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992
C:Genetics:
A:Gene: 84
A:Start codon: GTG

Query Match 34.3%; Score 55.5; DB 2; Length 66;
Best Local Similarity 43.8%; Pred. No. 0.52;
Matches 14; Conservative 2; Mismatches 7; Indels 9; Gaps 2;

OY 5 YGL-----RPGSSGPSLOYIKANSKFIGITEL 31
DB 36 YGFEVDWYEPGESG----YIKRNGFVGTWEV 63
|||||

RESULT 11

G48677
Ig heavy chain V-D-J region (419.1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C:Accession: G48677
R:Tassignon, J.; Brait, M.; Jamila, I.; Urbain, J.; Gottlieb, P.; Brown, A.; Hasemann Proc. Natl. Acad. Sci. U.S.A. 90, 9508-9512, 1993
A:Title: Molecular characterization of monoclonal CRI-A-positive anti-arsonate antibody
A:Reference number: A48677; MUID:94022404
A:Accession: G48677
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-123 <TAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 34.0%; Score 55; DB 2; Length 123;
Best Local Similarity 40.0%; Pred. No. 1.2;
Matches 16; Conservative 1; Mismatches 11; Indels 12; Gaps 2;

OY 4 SYGL-----RPGSS-----GPSLOQYKANSKFITEL 31
DB 31 SYGVNWKORPGQGLEWIGYINPGNDYIKYNEKFKGTTL 70
|||||

RESULT 12

RHAQ1
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
A:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.1;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
|||||

A:Gene: GnRH-I
C:Superfamily: gonadoliberin

Query Match 35.8%; Score 58; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||

Db 25 HWSYGLRPG 33

RESULT 6

RHMSG

gonadoliberin precursor - mouse
N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing
N:Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: A47578
R:Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolics,
Science 234, 1366-1371, 1986
A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo
A:Reference number: A47578; MUID:87069928
A:Accession: A47578
A:Molecule type: DNA
A:Residues: 1-90 <MAS>
A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
C:Genetics:
A:Introns: 45/3; 77/3
C:Function:

A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin #status predicted <GLB>
F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F:22/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predicted
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 35.8%; Score 58; DB 1; Length 90;

Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
|||||

Db 23 HWSYGLRPG 31

RESULT 7

RHHUG

gonadoliberin precursor [validated] - human
N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing
N:Contains: gonadoliberin-associated protein (GAP); progadoliberin
C:Species: Homo sapiens (man)
C:Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone ge
A:Reference number: S05308; MUID:89366682
A:Accession: S05308

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-92 <HAY>

A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956

R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.

Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986

A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
A:Reference number: A94090; MUID:86094338

A:Accession: A26173

A:Molecule type: mRNA

A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
A:Experimental source: hypothalamus
R:Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasin
A:Reference number: A93342; MUID:85012739
A:Accession: A93342
A:Molecule type: mRNA
A:Residues: 1-15,'S',17-92 <SEE>
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
A:Experimental source: placenta
R:Tan, L.; Rousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized
A:Reference number: A90108; MUID:83126573
A:Accession: A90108
A:Molecule type: protein
A:Residues: 24-33 <TAN>
A:Experimental source: placental trophoblasts
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amste
FEBS Lett. 346, 203-206, 1994
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by
A:Reference number: S45718; MUID:94283597
A:Contents: annotation; degradation pathway of synthetic hormone
C:Genetics:
A:Gene: GDB:GNRH; LHRH; GRH
A:Cross-references: GDB:I133746; OMIM:227200; OMIM:152760
A:Map position: 8p21-8p11.2
A:Introns: 47/3; 79/3
C:Function:

A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitrop
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activi
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progadoliberin #status predicted <PCN>
F:24-33/Product: gonadoliberin #status experimental <HAY>
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
F:22/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status experi
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 35.8%; Score 58; DB 1; Length 92;

Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10

|||||

Db 25 HWSYGLRPG 33

RESULT 8

RHRTG

gonadoliberin precursor - rat
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hor
N:Contains: gonadoliberin; prolactin release-inhibiting factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A40147; B26173; A48410
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic
A:Reference number: A40147; MUID:89384661
A:Accession: A40147

A:Molecule type: DNA

A:Residues: 1-92 <BON>

A:Cross-references: GB:M31670; NID:g204447; PIDN:AAA41264.1; PID:g204448

R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.

Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986

A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gona
A:Reference number: A94090; MUID:86094338

A:Accession: B26173

A:Molecule type: mRNA

R:de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A:Reference number: S69348; MUID:95262688
A:Accession: S69348
A:Molecule type: protein
A:Residues: 2-31 <DEF>
C:Comment: The source of this protein was an extrachromosomal plasmid.
C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
C:Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglio
C:Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
C:Function:
A:Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synapt
C:Superfamily: tetanus toxin
C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F:2-457/Product: tentoxylisin light chain (fragment A) #status predicted <TTL>
F:461-1315/Product: tentoxylisin heavy chain (fragment B.C) #status experimental <TTH>
F:461-864/Domain: channel forming (fragment B) #status predicted <TXB>
F:865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F:233,237/Binding site: zinc (His) #status predicted
F:234/Active site: Glu #status predicted

Query Match 46.9%; Score 76; DB 1; Length 1315;
Best Local Similarity 93.8%; Pred. No. 0.016;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYKANSKFIGITEL 31
:|||||
DB 829 MQYKANSKFIGITEL 844

RESULT 2

RHPGG
gonadoliblerin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase metho
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A:Note: the synthetic and natural
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117344
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C:Superfamily: gonadoliblerin
C:Keywords: amidated carboxyl end; hormone: hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2_HWSYGLRPG 10
:|||||
DB 2 HWSYGLRPG 10

RESULT 3

RHSHG
gonadoliblerin - sheep.
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780
A:Molecule type: protein
A:Residues: 1-10 <BUR>
A:Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliblerin
C:Keywords: amidated carboxyl end; hormone: hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
:|||||
DB 2 HWSYGLRPG 10

RESULT 4

I78541
gonadoliblerin precursor - rhesus macaque (fragment)
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I78541
R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A:Title: Developmental expression of the genes encoding transforming growth factor al
A:Reference number: I58134; MUID:95124501
A:Accession: I78541
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-67 <RES>
A:Cross-references: GB:S75918; NID:g912831; PIDN:AAB33096.1; PID:g912832
C:Superfamily: gonadoliblerin

Query Match 35.8%; Score 58; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
:|||||
DB 7 HWSYGLRPG 15

RESULT 5

I51423
gonadoliblerin precursor - African clawed frog
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I51423
R:Hayes, W.P.; Wray, S.; Battey, J.F.
Endocrinology 134, 1835-1845, 1994
A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved c
A:Reference number: I51423; MUID:94185563
A:Accession: I51423
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <HAY>
A:Cross-references: GB:L28040; NID:g496291; PIDN:AAA49728.1; PID:g496292
C:Genetics:

Query Match 35.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2_HWSYGLRPG 10
:|||||
DB 2 HWSYGLRPG 10

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 7.22009 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-15

Perfect score: 162

Sequence: 1 XHWSYGLRPGSSGSLQYIKANSKFIGITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	76	46.9	1315	1 BTCLTN	tentoxylisin (EC 3
2	58	35.8	10	1 RHPGG	gonadoliberin - pi
3	58	35.8	10	1 RHSHG	gonadoliberin - sh
4	58	35.8	67	2 I78541	gonadoliberin prec
5	58	35.8	89	2 I51423	gonadoliberin prec
6	58	35.8	90	1 RHMSG	gonadoliberin prec
7	58	35.8	92	1 RHHUG	gonadoliberin prec
8	58	35.8	92	1 RHRTG	gonadoliberin prec
9	56	34.6	719	2 T52510	gonadoliberin prec
10	55.5	34.3	66	2 S31029	hypothetical prote
11	55	34.0	123	2 G48677	gene 84 protein -
12	54	33.3	10	1 RHAQ1	Ig heavy chain V-D
13	54	33.3	92	2 I50644	gonadoliberin I -
14	53.5	33.0	374	2 E95361	gonadoliberin I pr
15	52	32.1	98	2 I50739	probable mucrone
16	52	32.1	102	2 PH1491	gonadotropin-rela
17	52	32.1	119	2 PH1518	Ig heavy chain V r
18	52	32.1	119	2 PH1519	Ig heavy chain V r
19	52	32.1	135	2 PH1494	Ig heavy chain V r
20	52	32.1	575	2 T06353	isocitrate lyase (
21	51	31.5	112	2 C27887	Ig kappa chain V r
22	51	31.5	115	2 S38715	Ig kappa chain V r
23	51	31.5	123	2 F48677	Ig heavy chain V-D
24	51	31.5	208	2 AG2249	hypothetical prote
25	50.5	31.2	1494	2 T14355	protein-tyrosine-p
26	50	30.9	80	1 RH1D1S	gonadoliberin I pr
27	50	30.9	120	2 A49043	Ig kappa chain V r
28	50	30.9	224	2 A53143	testis-determining
29	50	30.9	249	2 A14197	36K antigen pra -

RESULT 1

BTCLTN

tentoxylisin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin

C:Species: Clostridium tetani

C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999

C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,

EMBO J. 5, 2495-2502, 1986

A>Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b

A:Reference number: A25689; MUID:87053814

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770

R:Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A>Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; MUID:87040747

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774

A:Experimental source: strain CN3911

R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.

J. Bacteriol. 165, 21-27, 1986

A>Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C i

A:Reference number: A25194; MUID:86085872

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <FA2>

A:Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921

A:Accession: B25194

A:Molecule type: protein

A:Residues: 865-894 <FA3>

R:Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.

Infect. Immun. 57, 3588-3593, 1989

A>Title: Isolation, purification, and characterization of fragment B, the NH-2-termin

A:Reference number: A60759; MUID:90035436

A:Accession: A60759

A:Molecule type: protein

A:Residues: 461-475 <MAT>

R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.

J. Immunol. 142, 394-402, 1989

A>Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: JS0098; MUID:89093918

A:Contents: annotation; epitope region

R:Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,

Nature 359, 832-835, 1992

A>Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo

A:Reference number: S27125; MUID:93063293

A:Contents: annotation


```
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 15 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
; PCT-US93-11703-69
;
; Query Match      45.7%; Score 74; DB 5; Length 15;
; Best Local Similarity 100.0%; Pred. No. 7e-05;
; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 17 QYIKANSKFIGITEL 31
; DB 1 QYIKANSKFIGITEL 15
```

Search completed: October 10, 2002, 16:14:04
Job time : 6.55769 secs

```
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gollrick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-460-502-7

Query Match 45.7%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 QYIKANSKFGITEL 31
Db 1 QYIKANSKFGITEL 15

RESULT 13
US-09-046-373-2
; Sequence 2, Application US/09046373
; Patent No. 6235714
; GENERAL INFORMATION:
; APPLICANT: Sudhir Paul
; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Gololobov
; TITLE OF INVENTION: Methods for Identifying Inducers and
; TITLE OF INVENTION: Inhibitors of Catalytic Antibodies, Compositions and Their
; TITLE OF INVENTION: Use
; FILE REFERENCE: UNMC 63123
; CURRENT APPLICATION NUMBER: US/09/046,373
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
; US-09-046-373-2

Query Match 45.7%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 QYIKANSKFGITEL 31
Db 1 QYIKANSKFGITEL 15

RESULT 14
US-09-188-082-6
; Sequence 6, Application US/09188082
; Patent No. 6270765
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezia Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
```

```
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,082
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/661,052
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: Internal
; US-09-188-082-6

Query Match 45.7%; Score 74; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 QYIKANSKFGITEL 31
Db 1 QYIKANSKFGITEL 15

RESULT 15
PCT-US93-11703-69
; Sequence 69, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
```

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-35

Query Match 46.9%; Score 76; DB 2; Length 47;
Best Local Similarity 84.2%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GPSLOVYIKANSKFIGITEL 31
DB 17 GKKQYIKANSKFIGITEL 35

RESULT 10

US-08-319-704-10
Sequence 10, Application US/08319704

Patent No. 5814617

GENERAL INFORMATION:

APPLICANT: Hoffman, Stephen L.
APPLICANT: Charoenvit, Yupin
APPLICANT: Hedstrom, Richard C.
APPLICANT: Doolan, Denise L.

TITLE OF INVENTION: Protective 17 kDa Malaria Hepatic and
TITLE OF INVENTION: Erythrocytic stage immunogen and gene
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Naval Medical R & D Command
STREET: Bldg 1, T-12, 8901 Wisconsin Avenue

CITY: Bethesda

STATE: Maryland

COUNTRY: U.S.A

ZIP: 20889-5606

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/319,704

FILING DATE: 07-OCT-1994

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: A. David Spevack

REGISTRATION NUMBER: 24,743

REFERENCE/DOCKET NUMBER: 75,206

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 295-6759

TELEFAX: (301) 295-1022

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-319-704-10

Query Match 45.7%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
DB 1 QYIKANSKFIGITEL 15

RESULT 11

US-08-661-052-6

Sequence 6, Application US/08661052

Patent No. 5637243

GENERAL INFORMATION:
APPLICANT: Yashwant M. Deo
APPLICANT: Joel Goldstein
APPLICANT: Robert Graziano
APPLICANT: Chezian Somasundaram
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/661,052

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/484,172

FILING DATE: 07-JUNE-1995

ATTORNEY/AGENT INFORMATION:

NAME: Arnold, Beth E.

REGISTRATION NUMBER: 35,430

REFERENCE/DOCKET NUMBER: MXI-043CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-661-052-6

Query Match 45.7%; Score 74; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QYIKANSKFIGITEL 31
DB 1 QYIKANSKFIGITEL 15

RESULT 12

US-08-460-502-7

Sequence 7, Application US/08460502

Patent No. 5843464

GENERAL INFORMATION:

APPLICANT: Bakaletz, Lauren O.

APPLICANT: Kaumaya, Parvin T.

TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Calfee, Halter and Griswold

STREET: 800 Superior Avenue

CITY: Cleveland

STATE: Ohio

COUNTRY: U.S.A.

ZIP: 44114-2688

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:


```
;
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-57

Query Match 46.9%; Score 76; DB 1; Length 37;
Best Local Similarity 84.2%; Pred. No. 9.8e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GPSLOYIKANSKFIGITEL 31
Db 17 GKKQYIKANSKFIGITEL 35

RESULT 5
US-08-446-692-63
; Sequence 63, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488.351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446.692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-57

Query Match 46.9%; Score 76; DB 2; Length 37;
Best Local Similarity 84.2%; Pred. No. 9.8e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GPSLOYIKANSKFIGITEL 31
Db 17 GKKQYIKANSKFIGITEL 35

RESULT 7
US-08-488-351A-63
; Sequence 63, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
```


Query Match 51.9%; Score 84; DB 5; Length 24;
Best Local Similarity 77.3%; Pred. No. 3.8e-06;
Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 PGSSGPSLQYIKANSKFGITEL 30
Db 3 PSDFPFSVOYIKANSKFGITEL 24

RESULT 2
PCT-US95-08596-23
; Sequence 23, Application PC/TUS9508596
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Proinsulin Peptide Compounds for Detecting
; TITLE OF INVENTION: and Treating Type I Diabetes
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,220
; FILING DATE: 08-JULY-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-092PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
PCT-US95-08596-23

Query Match 46.9%; Score 76; DB 5; Length 17;
Best Local Similarity 93.8%; Pred. No. 4e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LQYIKANSKFGITEL 31
Db 2 MOYIKANSKFGITEL 17

RESULT 3
PCT-US93-11703-63
; Sequence 63, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville

; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-11703-63

Query Match 46.9%; Score 76; DB 5; Length 31;
Best Local Similarity 93.8%; Pred. No. 8e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LQYIKANSKFGITEL 31
Db 9 MOYIKANSKFGITEL 24

RESULT 4
US-08-446-692-57
; Sequence 57, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 6.55769 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-15
Perfect score: 162
Sequence: 1 XHWSYGLRPGSSGSLQYIKANSKFITEL 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	51.9	24	5	PCT-US92-07218-31
2	76	46.9	17	5	PCT-US95-08596-23
3	76	46.9	31	5	PCT-US93-11703-63
4	76	46.9	37	1	US-08-446-692-57
5	76	46.9	37	1	US-08-446-692-63
6	76	46.9	37	2	US-08-488-351A-57
7	76	46.9	37	2	US-08-488-351A-63
8	76	46.9	47	1	US-08-446-692-35
9	76	46.9	47	2	US-08-488-351A-35
10	74	45.7	15	2	US-08-319-704-10
11	74	45.7	15	2	US-08-661-052-6
12	74	45.7	15	2	US-08-460-502-7
13	74	45.7	15	4	US-09-046-373-2
14	74	45.7	15	4	US-09-188-082-6
15	74	45.7	15	5	PCT-US93-11703-69
16	74	45.7	16	4	US-09-248-588-55
17	74	45.7	17	1	US-08-446-692-4
18	74	45.7	17	2	US-08-488-351A-4
19	74	45.7	17	3	US-09-100-409A-40
20	74	45.7	17	5	PCT-US95-13841-7
21	74	45.7	27	1	US-08-446-692-13
22	74	45.7	27	2	US-08-488-351A-13
23	73	45.1	32	1	US-08-186-266-9
24	72	44.4	27	5	PCT-US92-07218-32
25	72	44.4	29	3	US-09-075-257A-13
26	72	44.4	29	3	US-09-075-257A-14
27	72	44.4	29	4	US-09-534-639-13

28	72	44.4	29	4	US-09-534-639-14	Sequence 14, Appl
29	72	44.4	30	5	PCT-US92-07218-29	Sequence 29, Appl
30	72	44.4	50	4	US-09-171-969-7	Sequence 7, Appl
31	71	43.8	17	1	US-08-188-223-6	Sequence 6, Appl
32	71	43.8	17	4	US-08-968-466-6	Sequence 6, Appl
33	71	43.8	17	4	US-08-478-546B-6	Sequence 6, Appl
34	70	43.2	14	1	US-08-186-266-5	Sequence 5, Appl
35	70	43.2	14	1	US-08-305-871A-5	Sequence 5, Appl
36	70	43.2	14	1	US-08-465-167A-18	Sequence 18, Appl
37	70	43.2	14	2	US-08-817-933A-9	Sequence 9, Appl
38	70	43.2	14	5	PCT-US92-07218-15	Sequence 15, Appl
39	70	43.2	14	5	PCT-US92-07218-30	Sequence 30, Appl
40	70	43.2	14	5	PCT-US93-02121-95	Sequence 95, Appl
41	70	43.2	19	1	US-08-787-547-41	Sequence 41, Appl
42	70	43.2	24	5	PCT-US92-07218-25	Sequence 25, Appl
43	70	43.2	24	5	PCT-US95-02121-110	Sequence 110, App
44	70	43.2	27	5	PCT-US92-07218-26	Sequence 26, Appl
45	70	43.2	27	5	PCT-US92-07218-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
PCT-US92-07218-31
; Sequence 31, Application PC/TUS9207218
; GENERAL INFORMATION:
; APPLICANT: Vitiello, Maria A.
; APPLICANT: Chesnut, Robert W.
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: PCT/US92/07218
; FILING DATE: 19920826
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14137-26-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-07218-31

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/188,082
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/661,052
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold, Beth E.
 REGISTRATION NUMBER: 35,430
 REFERENCE//DOCKET NUMBER: MXI-043CP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)227-5941
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 21 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal
 US-09-188-082-12

Query Match 56.0%; Score 112; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPPKVSASHLE 37
 |||||
 Db 1 FNNFTVSFWLRVPPKVSASHLE 21

Search completed: October 10, 2002, 16:14:04
 Job time : 7.82692 secs

ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/089,595
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,774
FILING DATE: 03-October-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6153728man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: amino acid
TOPOLOGY: linear
US-09-089-595-5

Query Match 56.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFMLRVPKVSASHLE 37
DB 1 FNNFTVSFMLRVPKVSASHLE 21
|||||

RESULT 13
US-09-382-855-5
Sequence 5, Application US/09382855
Patent No. 6174692
GENERAL INFORMATION:
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie, Pierre;
APPLICANT: Carrotlini, Jean-Charles; Carrel, Stefan; Reed, Daryl
TITLE OF INVENTION: MAGE-10 ENCODING CDNA, The Tumor Rejection
TITLE OF INVENTION: Antigen Precursors MAGE-10, Antibodies Specific To The Molecule,
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/382,855
FILING DATE: 25-August-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/089,595
FILING DATE: 02-June-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,774
FILING DATE: 03-October-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Schofield, Mary Anne
REGISTRATION NUMBER: 36,669
REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/MAS
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: amino acid
TOPOLOGY: linear
US-09-382-855-5

Query Match 56.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFMLRVPKVSASHLE 37
DB 1 FNNFTVSFMLRVPKVSASHLE 21
|||||

RESULT 14
US-09-183-714B-5
Sequence 5, Application US/09183714B
Patent No. 6221593
GENERAL INFORMATION:
APPLICANT: Boon-Falieu, Thierry
APPLICANT: Brasseur, Francis
APPLICANT: Rimoldi, Donata
APPLICANT: De Plaen, Etienne
TITLE OF INVENTION: Method for Determining Cancer by Determining Expression
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/183,714B
CURRENT FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: US 08/724,774
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 7
SEQ ID NO 5
LENGTH: 21
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
US-09-183-714B-5

Query Match 56.0%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFMLRVPKVSASHLE 37
DB 1 FNNFTVSFMLRVPKVSASHLE 21
|||||

RESULT 15
US-09-188-082-12
Sequence 12, Application US/09188082
Patent No. 6270765
GENERAL INFORMATION:
APPLICANT: Yashwant M. Deco
APPLICANT: Joel Goldstein
APPLICANT: Robert Graziano
APPLICANT: Chezia Somasundaram
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MXI-043CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-661-052-12

Query Match          56.0%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSEFWLRPKVSASHLE 37
DB 1 FNNFTVSEFWLRPKVSASHLE 21

RESULT 10
US-08-460-502-8
Sequence 8, Application US/08460502
Patent No. 5843464
GENERAL INFORMATION:
APPLICANT: Bakaletz, Lauren O.
APPLICANT: Kaumaya, Parvati T.
TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter and Griswold
STREET: 800 Superior Avenue
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,502
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Golrick, Mary E.
REGISTRATION NUMBER: 34,829
REFERENCE/DOCKET NUMBER: 22727/00120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8458
TELEFAX: (216) 241-0816
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-502-8

Query Match          56.0%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSEFWLRPKVSASHLE 37
DB 1 FNNFTVSEFWLRPKVSASHLE 21
```

```
RESULT 11
US-08-724-774B-5
Sequence 5, Application US/08724774B
Patent No. 5908778
GENERAL INFORMATION:
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie,
APPLICANT: Pierre; Cerrotilini, Jean-Charles; Carrel,
APPLICANT: Stefan; Reed, Daryl
TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
TITLE OF INVENTION: Rejection Antigen Precuros MAGE-10,
TITLE OF INVENTION: Antibodies Specific To The Molecule, and
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,774B
FILING DATE: 03-October-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5908778man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: amino acid
TOPOLOGY: linear
US-08-724-774B-5

Query Match          56.0%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSEFWLRPKVSASHLE 37
DB 1 FNNFTVSEFWLRPKVSASHLE 21

RESULT 12
US-09-089-595-5
Sequence 5, Application US/09089595
Patent No. 6153728
GENERAL INFORMATION:
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie,
APPLICANT: Pierre; Cerrotilini, Jean-Charles; Carrel,
APPLICANT: Stefan; Reed, Daryl
TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
TITLE OF INVENTION: Rejection Antigen Precuros MAGE-10,
TITLE OF INVENTION: Antibodies Specific To The Molecule, and
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
```

Best Local Similarity 95.5%; Pred. No. 4.1e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSWLRVPKVSASHLE 37
:|||||
Db 83 MFNFTVSWLRVPKVSASHLE 104

RESULT 7

US-08-668-381A-5
; Sequence 5, Application US/08668381A
; Patent No. 5780024
; GENERAL INFORMATION:
; APPLICANT: Brown, Robert H.
; APPLICANT: Fishman, Paul S.
; APPLICANT: Francis, Jonathan W.
; APPLICANT: Hosler, Betsy A.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/668,381A
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,473
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,164
; REFERENCE/DOCKET NUMBER: 00786/269001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-668-381A-5

Query Match 57.0%; Score 114; DB 1; Length 618;
Best Local Similarity 95.5%; Pred. No. 6e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNFTVSWLRVPKVSASHLE 37
:|||||
Db 249 MFNFTVSWLRVPKVSASHLE 270

RESULT 8

US-07-610-525-1
; Sequence 1, Application US/07610525
; Patent No. 5196512
; GENERAL INFORMATION:
; APPLICANT: BIANCHI Elisabetta
; APPLICANT: PESSI Antonello
; APPLICANT: CORRADIN Giampaolo
; TITLE OF INVENTION: SYNTHETIC PEPTIDES USEFUL AS

; TITLE OF INVENTION: UNIVERSAL CARRIERS FOR THE PREPARATION OF IMMUNOGENIC CONJU
; TITLE OF INVENTION: AND THEIR USE IN THE DEVELOPMENT OF SYNTHETIC VACCINES.
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHEA & GOULD
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10020-1193

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/610,525
FILING DATE: 19901108
CLASSIFICATION: 424
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-827-3000
TELEFAX: 212-840-6702
TELEX: 423973
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acid residues
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: no
FRAGMENT TYPE: Internal fragment
US-07-610-525-1

Query Match 56.0%; Score 112; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSWLRVPKVSASHLE 37
:|||||
Db 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 9

US-08-661-052-12
; Sequence 12, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISD
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995

APPLICATION NUMBER: PCT/GB90/00943
FILING DATE: 20-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8914122.0
FILING DATE: 20 June 1989
ATTORNEY/AGENT INFORMATION:
NAME: Mary J. Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 117-134
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-110-786a-8

Query Match 57.0%; Score 114; DB 1; Length 452;
Best local Similarity 95.5%; Pred. No. 4.1e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSEFLRPKVSASHLE 37
:|||||
DB 83 MFNNFTVSEFLRPKVSASHLE 104

RESULT 5
US-08-280-228-2
Sequence 2, Application US/08280228
Patent No. 5571694
GENERAL INFORMATION:
APPLICANT: Makoff Dr, Andrew J
APPLICANT: Romanos Dr, Michael A
APPLICANT: Clare Dr, Jeffrey J
APPLICANT: Fairweather Dr, Neil F
TITLE OF INVENTION: VACCINES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE P.C.
STREET: 1100 No. 5571694th Glebe Road
CITY: Arlington,
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/280,228
FILING DATE: 25-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/618,312
FILING DATE: 27-NOV-1990
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8926832.0
FILING DATE: 28-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9006097.1
FILING DATE: 17-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mary J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 117-163
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000

TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-280-228-2

Query Match 57.0%; Score 114; DB 1; Length 452;
Best local Similarity 95.5%; Pred. No. 4.1e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSEFLRPKVSASHLE 37
:|||||
DB 83 MFNNFTVSEFLRPKVSASHLE 104

RESULT 6
US-08-280-228-4
Sequence 4, Application US/08280228
Patent No. 5571694
GENERAL INFORMATION:
APPLICANT: Makoff Dr, Andrew J
APPLICANT: Romanos Dr, Michael A
APPLICANT: Clare Dr, Jeffrey J
APPLICANT: Fairweather Dr, Neil F
TITLE OF INVENTION: VACCINES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE P.C.
STREET: 1100 No. 5571694th Glebe Road
CITY: Arlington,
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/280,228
FILING DATE: 25-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/618,312
FILING DATE: 27-NOV-1990
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8926832.0
FILING DATE: 28-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9006097.1
FILING DATE: 17-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mary J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 117-163
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-280-228-4
Query Match 57.0%; Score 114; DB 1; Length 452;


```
RESULT 2
US-07-618-312A-2
; Sequence 2, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia,
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-618-312A-2

Query Match          57.0%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 4.1e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-618-312A-4

Query Match          57.0%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 4.1e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 16 LFNNFTVSFWLVRPKVSASHLE 37
      :|:|||||
DB 83 MFNNFTVSFWLVRPKVSASHLE 104

RESULT 4
US-08-110-786A-8
; Sequence 8, Application US/08110786A
; Patent No. 5443966
; GENERAL INFORMATION:
; APPLICANT: FAIRWEATHER, Neil Fraser
; APPLICANT: MAKOFF, Andrew Joseph
; TITLE OF INVENTION: Expression of tetanus toxin fragment C
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vandertye P.C.
; STREET: 1100 No. 5443966th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,786A
; FILING DATE: 23-AUG-1993 1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,337
; FILING DATE: 29-NOV-1991
; PRIOR APPLICATION DATA:
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 7.82692 Seconds

(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-14

Perfect score: 200

Sequence: 1 XHMSYGLRPGSSGSLFNFTYSPMLRPVKVASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	57.0	31	5 PCT-US93-11703-64	Sequence 64, Appl
2	114	57.0	452	1 US-07-618-312A-2	Sequence 2, Appl
3	114	57.0	452	1 US-07-618-312A-4	Sequence 4, Appl
4	114	57.0	452	1 US-08-110-786A-8	Sequence 8, Appl
5	114	57.0	452	1 US-08-280-228-2	Sequence 2, Appl
6	114	57.0	452	1 US-08-280-228-4	Sequence 4, Appl
7	114	57.0	618	1 US-08-668-381A-5	Sequence 5, Appl
8	112	56.0	21	1 US-07-610-525-1	Sequence 1, Appl
9	112	56.0	21	2 US-08-661-052-12	Sequence 12, Appl
10	112	56.0	21	2 US-08-460-502-8	Sequence 8, Appl
11	112	56.0	21	2 US-08-724-774B-5	Sequence 5, Appl
12	112	56.0	21	4 US-09-089-595-5	Sequence 5, Appl
13	112	56.0	21	4 US-09-382-855-5	Sequence 5, Appl
14	112	56.0	21	4 US-09-183-714B-5	Sequence 5, Appl
15	112	56.0	21	4 US-09-188-082-12	Sequence 12, Appl
16	112	56.0	21	4 US-09-171-969-10	Sequence 10, Appl
17	112	56.0	21	5 PCT-US93-11703-66	Sequence 66, Appl
18	112	56.0	32	2 US-08-446-692-14	Sequence 14, Appl
19	112	56.0	32	2 US-08-488-351A-14	Sequence 14, Appl
20	107	53.5	22	1 US-08-446-692-5	Sequence 5, Appl
21	107	53.5	22	2 US-08-488-351A-5	Sequence 5, Appl
22	107	53.5	22	3 US-09-100-409A-41	Sequence 41, Appl
23	107	53.5	22	5 PCT-US93-13841-8	Sequence 8, Appl
24	100	50.0	19	1 US-07-610-525-2	Sequence 2, Appl
25	94.5	47.2	20	1 US-08-319-704-11	Sequence 11, Appl
26	88	44.0	17	1 US-07-610-525-3	Sequence 3, Appl
27	87.5	43.8	22	2 US-08-817-933A-8	Sequence 8, Appl

ALIGNMENTS

28	79	39.5	14	1	US-07-610-525-6	Sequence 6, Appl
29	79	39.5	14	1	US-08-787-547-43	Sequence 43, Appl
30	79	39.5	15	1	US-07-610-525-4	Sequence 4, Appl
31	79	39.5	15	1	US-08-787-547-44	Sequence 44, Appl
32	71	35.5	17	1	US-08-188-223-6	Sequence 6, Appl
33	71	35.5	17	4	US-08-968-466-6	Sequence 6, Appl
34	71	35.5	17	4	US-08-478-546B-6	Sequence 6, Appl
35	69	34.5	13	1	US-07-610-525-5	Sequence 5, Appl
36	66	33.0	699	2	US-08-694-865-16	Sequence 16, Appl
37	66	33.0	699	3	US-09-124-491-16	Sequence 16, Appl
38	65	32.5	438	1	US-08-480-604A-23	Sequence 23, Appl
39	65	32.5	438	2	US-08-405-496A-23	Sequence 23, Appl
40	65	32.5	438	4	US-08-915-136-23	Sequence 23, Appl
41	65	32.5	462	1	US-08-480-604A-26	Sequence 26, Appl
42	65	32.5	462	2	US-08-405-496A-26	Sequence 26, Appl
43	65	32.5	462	4	US-08-915-136-26	Sequence 26, Appl
44	65	32.5	1296	1	US-08-480-604A-28	Sequence 28, Appl
45	65	32.5	1296	2	US-08-405-496A-28	Sequence 28, Appl

RESULT 1
PCT-US93-11703-64
Sequence 64, Application PC/TUS9311703
GENERAL INFORMATION:
APPLICANT: Chiron Mimotopes Pty. Ltd.
TITLE OF INVENTION: T-Cell Epitopes
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Grant D. Green
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.308
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11703
FILING DATE: 28-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,852
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0222.101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-11703-64
Query Match 57.0%; Score 114; DB 5; Length 31;
Best Local Similarity 95.3%; Pred. No. 1.6e-10;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 16 MFNFTVSPMLRPVKVASHLE 37
Db 6 MFNFTVSPMLRPVKVASHLE 27

XX 28-MAY-1999; 99US-0322289.
 XX (NEUR-) NEURALAB LTD.
 XX Schenk DB, Bard F, Vasquez NJ, Yednock T;
 XX WPI; 2001-032104/04.
 XX Preventing or treating a disease associated with amyloid deposits,
 PT especially Alzheimer's disease, comprises administering amyloid
 PT specific antibody
 XX
 PS Disclosure; Page 31; 143pp; English.
 XX
 CC This invention describes a novel method of preventing or treating a
 CC disease associated with amyloid deposits of amyloid precursor protein
 CC (APP) Abeta fragments in the brain of a patient, which comprises
 CC administering to the patient: (a) an antibody that binds to Abeta, the
 CC antibody binds to an amyloid deposit and induces a clearing response (Fc
 CC receptor mediated phagocytosis) against it (b) a polypeptide containing
 CC an N-terminal segment of at least residues 1-5 of Abeta; or (c) an agent
 CC that induces an immunogenic response against residues 1-3 to 7-11 of
 CC Abeta. The products of the invention have neurotropic and neuroprotective
 CC activity. The method is also useful for monitoring a course of treatment
 CC being administered to a patient e.g. active and passive immunization. The
 CC methods are useful for prophylactic and therapeutic treatment of
 CC Alzheimer's disease.
 XX
 SQ Sequence 43 AA;

Query Match 58.0%; Score 116; DB 22; Length 43;
 Best Local Similarity 100.0%; Pred. No. 8.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLVRPKVSASHLE 37
 DB 22 LFNNFTVSFWLVRPKVSASHLE 43

Search completed: October 10, 2002, 16:05:13
 Job time : 20.4487 secs

```

XX PF 20-JUL-2000; 2000WO-DK00413.
XX XX 20-JUL-1999; 99DK-0001014.
XX PR 26-JUL-1999; 99US-0145275.
XX XX (MEBI-) M & E BIOTECH AS.
XX PA Halkier T, Mouritsen S, Klysner S;
XX PI WPI; 2001-112680/12.
XX DR
XX XX
XX PT Increasing the muscle mass of animals used in meat production by down
XX PT regulating growth differentiation factor 8 (GDF-8) activity in the
XX PT animal through induction of anti-GDF-8 antibody production .
XX PS Example 1; Page 107-108; 110pp; English.
XX XX
XX CC The present sequence is that of AutoVac construct GDF-8 ext,
XX CC which consists of the C-terminal 160 amino acids of human growth
XX CC differentiation factor 8 (GDF-8, see AAF20131) with residues 16-36
XX CC substituted by the promiscuous tetanus toxin T-cell epitope P30 (see
XX CC AAB20144) and residues 37-51 substituted by tetanus toxin T-cell
XX CC epitope P2 (see AAB20143). It is an object of the invention to
XX CC produce a recombinant therapeutic vaccine that is capable of effecting
XX CC down-regulation of GDF-8 in order to increase the muscle growth
XX CC rate of farm animals. The vaccines (see AAB20145-53) are capable
XX CC of breaking autotolerance against autologous GDF-8. They comprise
XX CC the C-terminal portion of human GDF-8 in which a portion of the
XX CC native sequence is replaced by a T-cell epitope such as P30, with
XX CC minimal disturbance of the authentic 3-dimensional structure of
XX CC the protein. Nucleic acids encoding the GDF-8 variants can be used
XX CC for genetic immunisation of the animals. Down-regulation of GDF-8
XX CC activity can increase muscle mass by up to at least 45% in cattle,
XX CC pigs and poultry used for meat production, reducing the need for
XX CC antibiotic feed-additives. Anti-GDF8 vaccines can be used to
XX CC treat human diseases such as cancer cachexia where muscle atrophy is
XX CC pronounced and for patients suffering from acute and chronic heart
XX CC failure.
XX SQ Sequence 160 AA;

Query Match 58.5%; Score 117; DB 22; Length 160;
Best Local Similarity 95.7%; Pred. No. 2.9e-09; Indels 0; Gaps 0;
Matches 22; Conservative 1; Mismatches 0;

Qy 15 SLFNNFTVSFWLRVPKVSASHLE 37
:|||||
Db 14 ALFNNFTVSFWLRVPKVSASHLE 36

RESULT 14
AAY92653
ID AAY92653 standard; Peptide; 31 AA.
XX AC AAY92653;
XX XX
XX DT 10-AUG-2000 (first entry)
XX DE
XX DE PSMep010 - P30 inserted in hPSM insertion position 6.
XX KW Foreign epitope; P2; prostate specific membrane antigen; vaccination;
XX KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
XX KW prostate cancer; cell-associated peptide antigen.
XX XX Synthetic.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Peptide 6..26
XX FT /label= P30
XX XX
XX PN WO200020027-A2.

```

```

XX 13-APR-2000.
XX PD
XX PF 05-OCT-1999; 99WO-DK00525.
XX PR 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX XX (MEBI-) M & E BIOTECH AS.
XX XX
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX DR WPI; 2000-349917/30.
XX XX
XX PT Inducing immune responses to weakly immunogenic, tumor associated
XX PT peptide antigens for the treatment of breast and prostate cancer
XX PS Example 1; Page 117; 220pp; English.
XX XX
XX CC AAY92650-55 are peptides designed which correspond to the P2 and P30
XX CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
XX CC amino acids in each end. The flanking amino acids correspond to the
XX CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g.
XX CC T cell proliferation assays, but also for ELISA or other in vitro
XX CC assays. The claims detail a method for inducing immune responses against
XX CC weakly immunogenic cell-associated peptide antigens (PA) such as those
XX CC associated with cancers (i.e. self-proteins), for example, hPSM,
XX CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX CC method comprises effecting simultaneous presentation by antigen
XX CC producing cells (APCs) of the animals immune system of: (1) at least 1
XX CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX CC B-cell group derived from the cell-associated PA; and (2) at least 1
XX CC first T helper cell group which is foreign to the animal. Analogues of
XX CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX CC part of all known and predicted CTL and B-cell epitopes of the
XX CC respective PA and including at least one foreign T helper epitope (e.g.
XX CC P2 and/or P30) are also claimed. The method is used to treat prostate,
XX CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and
XX CC Her2, respectively.
XX SQ Sequence 31 AA;

Query Match 58.0%; Score 116; DB 21; Length 31;
Best Local Similarity 100.0%; Pred. No. 6e-10; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
|||||
Db 5 LFNNFTVSFWLRVPKVSASHLE 26

RESULT 15
AAB46177
ID AAB46177 standard; peptide; 43 AA.
XX AC AAB46177;
XX XX
XX DT 04-APR-2001 (first entry)
XX DE
XX DE Tetanus toxoid 830-844 + 947-967 epitope AN90542.
XX KW Amyloid deposit; APP; Abeta; brain; human; clearing response; neurotropic;
XX KW FC receptor mediated phagocytosis; immunogenic response; neuroprotective;
XX KW amyloid precursor protein; Alzheimer's disease.
XX XX
XX OS Clostridium tetani.
XX XX
XX PN WO200072880-A2.
XX XX
XX PD 07-DEC-2000.
XX XX
XX PF 26-MAY-2000; 2000WO-US14810.

```

```

XX Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.
XX
XX Homo sapiens.
OS Clostridium tetani.
XX
XX WO2000065058-A1.
XX
XX 02-NOV-2000.
XX
XX 19-APR-2000; 2000WO-DK00205.
XX
XX 23-APR-1999; 99DK-0000552.
XX 06-MAY-1999; 99US-0132811.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Klysner S;
XX
XX WPI; 2000-672791/65.
XX
XX Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -
XX
XX Example 7; Page 124; 172pp; English.
XX
XX The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.
XX
XX Sequence 124 AA;
SQ
Query Match 58.5%; Score 117; DB 21; Length 124;
Best Local Similarity 95.7%; Pred. No. 2.2e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 15 SLFNNFTVSFWLRVPKVSASHLE 37
Db 30 TLFNNTVSFWLRVPKVSASHLE 52

RESULT 12
AAB45515
ID AAB45515 standard; Protein; 124 AA.
XX
XX AAB45515;
XX
XX 26-FEB-2001 (first entry)
XX
XX Modified human interleukin-5 SEQ ID NO: 30.
XX
XX Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
KW cancer; eosinophilia; vaccine; allergic rhinitis.
XX
XX Homo sapiens.
OS Clostridium tetani.
XX
XX WO2000065058-A1.
XX
XX 02-NOV-2000.
XX
XX 19-APR-2000; 2000WO-DK00205.
XX
XX 23-APR-1999; 99DK-0000552.
XX 06-MAY-1999; 99US-0132811.
XX

```

```

XX (MEBI-) M & E BIOTECH AS.
XX
XX Klysner S;
XX
XX WPI; 2000-672791/65.
XX N-PSDB; AAC68868.
XX
XX Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -
XX
XX Example 7; Page 141; 172pp; English.
XX
XX The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.
XX
XX Sequence 124 AA;
SQ
Query Match 58.5%; Score 117; DB 21; Length 124;
Best Local Similarity 95.7%; Pred. No. 2.2e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 15 SLFNNFTVSFWLRVPKVSASHLE 37
Db 30 TLFNNTVSFWLRVPKVSASHLE 52

RESULT 13
AAB20153
ID AAB20153 standard; Protein; 160 AA.
XX
XX AAB20153;
XX
XX 30-APR-2001 (first entry)
XX
XX Growth differentiation factor 8 AutoVac construct GDF-8 ext.
XX
XX Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
XX Chimeric - Homo sapiens.
OS Chimeric - Clostridium tetani.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Region 1..15 /note= "identical to residues 215-230 of human
FT GDF-8"
FT Region 16..36 /note= "tetanus toxoid p30 epitope"
FT Region 37..51 /note= "tetanus toxoid p2 epitope"
FT Region 52..160 /note= "identical to residues 267-375 of human
FT GDF-8"
FT Misc-difference 124 /note= "Cys-124 may be substituted by Ser to avoid
FT disulfide bond formation"
FT Misc-difference 141..142 /note= "optionally replaced by Glu-Gly"
XX
XX WO200105820-A2.
XX
XX 25-JAN-2001.
XX

```

```

OS XX Synthetic.
FH Key Location/Qualifiers
FT Peptide 631..651
FT /label= P30
XX
XX WO200020027-A2.
XX 13-APR-2000.
XX 05-OCT-1999; 99WO-DK00525.
XX 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX (MEBI-) M & E BIOTECH AS.
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92659-62 are mutant immunogenized murine prostate specific membrane
CC antigen (PSM) constructs, which contain a foreign epitope, P30. The
CC analogues can be used to study whether autotolerance to mouse PSM can be
CC broken in mice by immunisation and/or DNA vaccination against murine PSM
CC using murine PSM analogues. Immunogenic analogues of PSM can be used in
CC the claimed method as an autovaccine to induce a CTL response. The
CC method is used for inducing immune responses against weakly immunogenic
CC cell-associated peptide antigens (PA) such as those associated with
CC cancers (self-proteins), e.g. human PSM, heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC group derived from the PA and/or at least 1 B-cell group derived from
CC the cell-associated PA; and (2) at least 1 first T helper cell group
CC which is foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.
CC Note: This sequence was constructed from the murine PSM splice variant
CC (AAY92624), which appears on pages 210-213 of the specification.
XX
XX Sequence 703 AA;
XX
XX Query Match 59.0%; Score 118; DB 21; Length 703;
XX Best Local Similarity 82.1%; Pred. No. 1.2e-08;
XX Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
Qy 10 GSSGFSLFNNFTVSFWLRPKVKSASHLE 37
Db 624 GLGGRPFNNFTVSFWLRPKVKSASHLE 651
XX
RESULT 10
AAY92660
ID AAY92660 standard; Protein; 761 AA.
XX
XX AAY92660;
XX
XX 10-AUG-2000 (first entry)
XX
XX Mutant murine prostate specific membrane antigen construct, mPSMY.
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
XX vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW

```

```

KW XX prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Mus musculus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Peptide 689..709
FT /label= P30
XX
XX WO200020027-A2.
XX 13-APR-2000.
XX 05-OCT-1999; 99WO-DK00525.
XX 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX (MEBI-) M & E BIOTECH AS.
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92659-62 are mutant immunogenized murine prostate specific membrane
CC antigen (PSM) constructs, which contain a foreign epitope, P30. The
CC analogues can be used to study whether autotolerance to mouse PSM can be
CC broken in mice by immunisation and/or DNA vaccination against murine PSM
CC using murine PSM analogues. Immunogenic analogues of PSM can be used in
CC the claimed method as an autovaccine to induce a CTL response. The
CC method is used for inducing immune responses against weakly immunogenic
CC cell-associated peptide antigens (PA) such as those associated with
CC cancers (self-proteins), e.g. human PSM, heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC group derived from the PA and/or at least 1 B-cell group derived from
CC the cell-associated PA; and (2) at least 1 first T helper cell group
CC which is foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.
CC Note: This sequence was constructed from the wild type murine PSM
CC (AAY92623), which appears on pages 204-206 of the specification.
XX
XX Sequence 761 AA;
XX
XX Query Match 59.0%; Score 118; DB 21; Length 761;
XX Best Local Similarity 82.1%; Pred. No. 1.3e-08;
XX Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
Qy 10 GSSGFSLFNNFTVSFWLRPKVKSASHLE 37
Db 682 GLGGRPFNNFTVSFWLRPKVKSASHLE 709
XX
RESULT 11
AAB45496
ID AAB45496 standard; Protein; 124 AA.
XX
XX AAB45496;
XX
XX 26-FEB-2001 (first entry)
XX
XX Modified human interleukin-5 SEQ ID NO: 8.
DE

```

```

PF 15-APR-1998; 98WO-DK00157.
XX
PR 24-APR-1997; 97US-0044187.
PR 15-APR-1997; 97DK-0000418.
XX
PA (FERR ) FARM LAB FERRING AS.
XX
PI Dalum I, Elsnor H, Jensen MR, Mouritsen S;
XX
DR WPI; 1998-594561/50.
DR N-PSDB; AAV68421.
XX
XX Modified human tumour necrosis factor-alpha - comprises
PT predominant T cell epitope, useful in vaccines to treat or
PT prevent TNF-associated diseases, e.g. cancer
XX
XX Example 1; Page 74-75; 134pp; English.
XX
XX The present sequence represents a modified human tumour necrosis
CC factor-alpha (TNF-alpha) analogue. The analogues have no residual
CC TNF activity and are immunogenic in a large proportion of the human
CC population (by using promiscuous epitopes). The TNF-alpha analogue
CC is able to generate, in humans, neutralizing antibodies to wild-type
CC human TNF alpha, has at least one fragment of TNF substituted by a
CC peptide containing an immunodominant T-cell epitope, and at least one
CC TNF-alpha B-cell epitope. The substitution causes a significant change
CC in the amino acid sequence of any one of the strands in the front
CC beta-sheet, any of the connecting loops or any of the B', I or D strands
CC in the back beta-sheet. The TNF-alpha analogues are used as vaccines for
CC treatment or prevention of diseases associated with excessive release
CC or activity of TNF-alpha, e.g. rheumatoid arthritis, Crohn's disease,
CC ulcerative colitis, cancer of any sort, disseminated sclerosis, diabetes,
CC psoriasis, osteoporosis and asthma.
XX
SQ Sequence 158 AA;
Query Match 59.2%; Score 118.5; DB 19; Length 158;
Best Local Similarity 80.6%; Pred. No. 1.7e-09;
Matches 25; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Oy 7 LRPGSGPSLFNFTVSFWLRVPKVSASHLE 37
:| | |||||
Db 2 VRSSSRTPS-FNNFTVSFWLRVPKVSASHLE 31

RESULT 8
AAB20149
ID AAB20149 standard; Protein; 109 AA.
XX
AC AAB20149;
XX
DT 30-APR-2001 (first entry)
XX
DE Growth differentiation factor 8 AutoVac construct GDF-8 P30-2.
XX
KW Growth differentiation factor 8; GDF-8; mycostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Clostridium tetani.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..48
FT /note= "identical to residues 267-314 of human
FT GDF-8"
FT
FT Region 49..69
FT /note= "tetanus toxoid P2 epitope"
FT
FT Region 70..109
FT /note= "identical to residues 336-375 of human
FT GDF-8"
FT
FT Misc-difference 73

/Note= "Cys-73 may be substituted by Ser to avoid
disulfide bond formation"
Misc-difference 90..91
/Note= "optionally replaced by Glu-Gly"
WO200105820-A2.
XX
XX 25-JAN-2001.
XX
XX 20-JUL-2000; 2000WO-DK00413.
XX
XX 20-JUL-1999; 99DK-0001014.
XX 26-JUL-1999; 99US-0145275.
XX (MEBI-) M & E BIOTECH AS.
XX
XX Halkier T, Mouritsen S, Klysner S;
XX
XX WPI; 2001-112680/12.
XX
XX Increasing the muscle mass of animals used in meat production by down
PT regulating growth differentiation factor 8 (GDF-8) activity in the
PT animal through induction of anti-GDF-8 antibody production
XX
XX Example 1; Page 101-102; 110pp; English.
XX
XX The present sequence is that of AutoVac construct GDF-8 P30-2,
CC comprising the 109 C-terminal amino acid residues of human
CC growth differentiation factor 8 (GDF-8) in which residues 49-69 are
CC replaced by the promiscuous tetanus toxin T-cell epitope P30 (see
CC AAB20144). It is an object of the invention to produce a
CC recombinant therapeutic vaccine that is capable of effecting
CC down-regulation of GDF-8 in order to increase the muscle growth
CC rate of farm animals. The vaccines (see AAB20145-53) are capable
CC of breaking autotolerance against autologous GDF-8. They comprise
CC the C-terminal portion of human GDF-8 in which a portion of the
CC native sequence is replaced by a T-cell epitope such as P30, with
CC minimal disturbance of the authentic 3-dimensional structure of
CC the protein. Nucleic acids encoding the GDF-8 variants can be used
CC for genetic immunisation of the animals. Down-regulation of GDF-8
CC activity can increase muscle mass by up to at least 45% in cattle,
CC pigs and poultry used for meat production, reducing the need for
CC antibiotic feed-additives. Anti-GDF8 vaccines can be used to
CC treat human diseases such as cancer cachexia where muscle atrophy is
CC pronounced and for patients suffering from acute and chronic heart
CC failure.
XX
SQ Sequence 109 AA;
Query Match 59.0%; Score 118; DB 22; Length 109;
Best Local Similarity 61.0%; Pred. No. 1.3e-09;
Matches 25; Conservative 2; Mismatches 8; Indels 6; Gaps 1;

Oy 3 WSYGLRPG-----SSGPSLFNFTVSFWLRVPKVSASHLE 37
|:|:| |||||
Db 29 WDIIAPRYKANYCSECFNFTVSFWLRVPKVSASHLE 69

RESULT 9
AA92662
ID AA92662 standard; Protein; 703 AA.
XX
AC AA92662;
XX
XX 10-AUG-2000 (first entry)
XX
DE Mutant murine F5M splice variant construct, mPSM'Y.
XX
KW Prostate specific membrane antigen; splice variant; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Mus musculus.

```


XX
23

```
XX Chimeric - Homo sapiens.
OS Chimeric - Clostridium tetani.
XX Synthetic.
XX Key
FH Location/Qualifiers
FT 1..83 /note= "identical to residues 267-349 of human
FT GDF-8"
FT
FT Region
FT 84..104
FT /note= "tetanus toxoid P2 epitope"
FT 105..109
FT Region
FT /note= "identical to residues 371-375 of human
FT GDF-8"
FT
FT Misc-difference 73
FT /note= "Cys-73 may be substituted by Ser to avoid
FT disulfide bond formation"
FT
FT Misc-difference 90..91
FT /note= "optionally replaced by Glu-Gly"
FT
XX WO200105820-A2.
XX 25-JAN-2001.
XX
XX 20-JUL-2000; 2000WO-DK00413.
XX
XX 20-JUL-1999; 99DK-0001014.
XX 26-JUL-1999; 99US-0145275.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Halkier T, Mouritsen S, Klysner S;
XX
XX WPI; 2001-112680/12.
XX
XX Increasing the muscle mass of animals used in meat production by down
XX regulating growth differentiation factor 8 (GDF-8) activity in the
XX animal through induction of anti-GDF-8 antibody production -
XX
XX Example 1; Page 104; 110pp; English.
XX
XX The present sequence is that of AutoVac construct GDF-8 P30-3B,
XX comprising the 109 C-terminal amino acid residues of human
XX growth differentiation factor 8 (GDF-8) in which residues 84-104
XX are replaced by the promiscuous tetanus toxin T-cell epitope P30
XX (see AAB20144). It is an object of the invention to produce a
XX recombinant therapeutic vaccine that is capable of effecting
XX down-regulation of GDF-8 in order to increase the muscle growth
XX rate of farm animals. The vaccines (see AAB20145-53) are capable
XX of breaking autotolerance against autologous GDF-8. They comprise
XX the C-terminal portion of human GDF-8 in which a portion of the
XX native sequence is replaced by a T-cell epitope such as P30, with
XX minimal disturbance of the authentic 3-dimensional structure of
XX the protein. Nucleic acids encoding the GDF-8 variants can be used
XX for genetic immunisation of the animals. Down-regulation of GDF-8
XX activity can increase muscle mass by up to at least 45% in cattle,
XX pigs and poultry used for meat production, reducing the need for
XX antibiotic feed-additives. Anti-GDF8 vaccines can be used to
XX treat human diseases such as cancer cachexia where muscle atrophy is
XX pronounced and for patients suffering from acute and chronic heart
XX failure.
XX
XX Sequence 109 AA;
XX
XX Query Match 59.8%; Score 119.5; DB 22; Length 109;
XX Best Local Similarity 67.6%; Pred. No. 7.9e-10;
XX Matches 25; Conservative 1; Mismatches 2; Indels 9; Gaps 1;
Oy 10 GSSGPSL-----FNNFTVSEWLRVPKVSASHLE 37
DB 68 GSAGPCCTPTKMSPIFNFTVSEWLRVPKVSASHLE 104
|||||
|||||
```

```
RESULT 4
AA92644
ID AA92644 standard; Protein; 750 AA.
XX
XX AC AA92644;
XX
XX 10-AUG-2000 (first entry)
XX
XX Mutant human prostate specific membrane antigen construct, hPSM6.3.
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
XX vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
XX prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 448..462
XX /label= P2
XX /note= "foreign epitope"
XX 210..230
XX Peptide /label= P30
XX /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
XX peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -; 220pp; English.
XX
XX AA92627-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX P30). The immunogenic analogues of PSM can be used in the claimed method
XX as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX antibody binding regions and cysteine residues involved in disulfide
XX bonds are preserved in the immunogenized forms. The method is used for
XX inducing immune responses against weakly immunogenic cell-associated
XX peptide antigens (PA) such as those associated with cancers
XX (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX method comprises effecting simultaneous presentation by antigen producing
XX cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX B-cell group derived from the cell-associated PA; and (2) at least 1
XX first T helper cell group which is foreign to the animal. Analogues of
XX human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX part of all known and predicted CTL and B-cell epitopes of the respective
XX PA and including at least one foreign T helper epitope are also claimed.
XX The method is used to treat prostate, prostate/breast or breast cancer.
XX when the PA is human PSM, FGF8b and Her2, respectively.
XX Note: This sequence was constructed from the wild type human PSM
XX (AA92619), which appears on pages 184-187 of the specification.
XX
XX Sequence 750 AA;
XX
XX Query Match 59.8%; Score 119.5; DB 21; Length 750;
XX Best Local Similarity 78.1%; Pred. No. 7.5e-09;
```

```

XX PD 25-JAN-2001.
XX XX
XX PF 20-JUL-2000; 2000WO-DK00413.
XX XX
XX PR 20-JUL-1999; 99DK-0001014.
XX PR 26-JUL-1999; 99US-0145275.
XX XX
XX PA (MEBI-) M & E BIOTECH AS.
XX XX
XX PI Halkier T, Mouritsen S, Klysner S;
XX PI WPI; 2001-112680/12.
XX XX
XX PT Increasing the muscle mass of animals used in meat production by down
XX PT regulating growth differentiation factor 8 (GDF-8) activity in the
XX PT animal through induction of anti-GDF-8 antibody production
XX XX
XX PS Example 1; Page 102-103; 110pp; English.
XX XX
XX CC The present sequence is that of AutoVac construct GDF-8 P30-3A,
XX CC comprising the 109 C-terminal amino acid residues of human
XX CC growth differentiation factor 8 (GDF-8) in which residues 79-99 are
XX CC replaced by the promiscuous tetanus toxin T-cell epitope P30 (see
XX CC AAB20144). It is an object of the invention to produce a
XX CC recombinant therapeutic vaccine that is capable of effecting
XX CC down-regulation of GDF-8 in order to increase the muscle growth
XX CC rate of farm animals. The vaccines (see AAB20145-53) are capable
XX CC of breaking autotolerance against autologous GDF-8. They comprise
XX CC the C-terminal portion of human GDF-8 in which a portion of the
XX CC native sequence is replaced by a T-cell epitope such as P30, with
XX CC minimal disturbance of the authentic 3-dimensional structure of
XX CC the protein. Nucleic acids encoding the GDF-8 variants can be used
XX CC for genetic immunisation of the animals. Down-regulation of GDF-8
XX CC activity can increase muscle mass by up to at least 45% in cattle,
XX CC pigs and poultry used for meat production, reducing the need for
XX CC antibiotic feed-additives. Anti-GDF8 vaccines can be used to
XX CC treat human diseases such as cancer cachexia where muscle atrophy is
XX CC pronounced and for patients suffering from acute and chronic heart
XX CC failure.
XX XX
XX SQ Sequence 109 AA;
XX
XX Query Match 61.5%; Score 123; DB 22; Length 109;
XX Best Local Similarity 78.1%; Pred. No. 2.4e-10;
XX Matches 25; Conservative 2; Mismatches 1; Indels 4; Gaps 1;
XX
XX QY 10 GSSG----PSLFNFTVSFWLRVPKVSASHLE 37
XX ||||| | : ||||| ||||| ||||| |||||
XX Db 68 GSAGPCTPTFRNFTVSFWLRVPKVSASHLE 99
XX
XX RESULT 2
XX AAY92665
XX ID AAY92665 standard; Peptide; 216 AA.
XX AC AAY92665;
XX XX
XX DT 10-AUG-2000 (first entry)
XX XX
XX DE MUC-1 analogue containing foreign epitopes.
XX XX
XX KW Mucin repeat; MUC-1 analogue; vaccination; self-protein; cancer;
XX KW cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;
XX KW cell-associated peptide antigen; foreign epitope.
XX XX
XX OS Homo sapiens.
XX XX
XX FH Key : Location/Qualifiers
XX FT Peptide 61..75
XX FT Peptide /label= P2
XX FT Peptide 136..156
XX FT Peptide /label= P30

```

```

FT XX /note= "q"
XX PN WO200020027-A2.
XX PD 13-APR-2000.
XX XX
XX PF 05-OCT-1999; 99WO-DK00525.
XX XX
XX PR 05-OCT-1998; 98DK-0001261.
XX PR 20-OCT-1998; 98US-0105011.
XX XX
XX PA (MEBI-) M & E BIOTECH AS.
XX XX
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX XX
XX DR WPI; 2000-349917/30.
XX XX
XX PT Inducing immune responses to weakly immunogenic, tumor associated
XX PT peptide antigens for the treatment of breast and prostate cancer
XX XX
XX PS Example 4; Page -; 220pp; English.
XX XX
XX CC This is an immunogenized MUC-1 analogue containing foreign epitopes P2
XX CC and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate
XX CC specific membrane antigen (hPSM) can be used in the claimed method as an
XX CC autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX CC antibody binding regions and cysteine residues involved in disulfide
XX CC bonds are preserved in the immunogenized forms (see features table). 10
XX CC regions suitable for the insertion of foreign T helper epitopes were
XX CC identified. The method is used for inducing immune responses against
XX CC weakly immunogenic cell-associated peptide antigens (PA) such as those
XX CC associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)
XX CC and/or fibroblast growth factor 8b (FGF8b). The method comprises
XX CC effecting simultaneous presentation by antigen producing cells (APCs) of
XX CC the animals immune system of: (1) at least 1 CTL (cytotoxic
XX CC T-lymphocyte) group derived from the PA and/or at least 1 B-cell group
XX CC derived from the cell-associated PA; and (2) at least 1 first T helper
XX CC cell group which is foreign to the animal. Analogues of human PSM, human
XX CC Her2 and human/murine FGF8b comprising a substantial part of all known
XX CC and predicted CTL and B-cell epitopes of the respective PA and including
XX CC at least one foreign T helper epitope are also claimed. The method is
XX CC used to treat prostate, prostate/breast or breast cancer when the PA is
XX CC human PSM, FGF8b and Her2, respectively.
XX CC Note: This sequence does not appear in the specification. It was made
XX CC using the mucin repeat sequence (AAY92664), P2 and P30 (AAY92625-26),
XX CC which appear on pages 220, 213 and 214 respectively, of the
XX CC specification.
XX XX
XX SQ Sequence 216 AA;
XX
XX Query Match 60.5%; Score 121; DB 21; Length 216;
XX Best Local Similarity 64.1%; Pred. No. 1.1e-09;
XX Matches 25; Conservative 2; Mismatches 2; Indels 10; Gaps 1;
XX
XX QY 9 PGSSGP-----SLFNFTVSFWLRVPKVSASHLE 37
XX ||||| | : ||||| ||||| ||||| |||||
XX Db 118 PGSTAPPAGVTSAPDTRFNFTVSFWLRVPKVSASHLE 156
XX
XX RESULT 3
XX AAB20151
XX ID AAB20151 standard; Protein; 109 AA.
XX AC AAB20151;
XX XX
XX DT 30-APR-2001 (first entry)
XX XX
XX DE Growth differentiation factor 8 AutoVac construct GDF-8 P30-3B.
XX XX
XX KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
XX KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
XX KW cardiac; human; mutant; mutein.

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 19.4487 Seconds
(without alignments)
211.311 Million cell updates/sec

Title: US-09-848-834A-14
Perfect score: 200
Sequence: 1 XHWSYGLRPGSGSLFNNFTVSFWLRVVKVASHLE 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	123	61.5	22 AAB20150	Growth differentia
2	121	60.5	21 AAY92665	MUC-1 analogue con
3	119.5	59.8	20 AAB20151	Growth differentia
4	119.5	59.8	21 AAY92644	Mutant human prost
5	119	59.5	21 AAB45507	Modified murine in
6	119	59.5	122 AAB45524	Modified murine in
7	118.5	59.2	19 AAW81332	TNF30-1, a TNF- α
8	118	59.0	22 AAB20149	Growth differentia
9	118	59.0	703 AAY92662	Mutant murine PSM
10	118	59.0	21 AAY92660	Mutant murine pros
11	117	58.5	124 AAB45496	Modified human int

12	117	58.5	124	21	AAB45515	Modified human int
13	117	58.5	160	22	AAB20153	Growth differentia
14	116	58.0	31	21	AAY92653	PSMep010 - P30 in
15	116	58.0	43	22	AAB46177	Tetanus toxoid 830
16	116	58.0	43	22	AAB49076	Amyloid beta/tetan
17	116	58.0	72	22	AAB46190	Tetanus toxoid epi
18	116	58.0	109	22	AAB20148	Growth differentia
19	116	58.0	136	22	AAB49089	Amyloid beta tetan
20	116	58.0	145	21	AAB45530	Modified murine in
21	116	58.0	147	21	AAB45522	Modified human int
22	116	58.0	158	19	AAW81334	TNF30-3, a TNF- α
23	116	58.0	188	21	AAY84423	An osteoprotegerin
24	116	58.0	254	22	AAB20152	Growth differentia
25	116	58.0	750	21	AAY92627	Mutant human prost
26	116	58.0	750	21	AAY92628	Mutant human prost
27	116	58.0	750	21	AAY92629	Mutant human prost
28	116	58.0	750	21	AAY92630	Mutant human prost
29	116	58.0	750	21	AAY92631	Mutant human prost
30	116	58.0	750	21	AAY92637	Mutant human prost
31	116	58.0	750	21	AAY92638	Mutant human prost
32	116	58.0	750	21	AAY92639	Mutant human prost
33	116	58.0	750	21	AAY92642	Mutant human prost
34	115	57.5	693	21	AAY92647	Mutant human PSM a
35	115	57.5	693	21	AAY92648	Mutant human PSM a
36	115	57.5	750	21	AAY92645	Mutant human prost
37	115	57.5	750	21	AAY92646	Mutant human prost
38	114	57.0	130	21	AAB45497	Modified human int
39	114	57.0	130	21	AAB45516	Modified human int
40	114	57.0	452	12	AA12471	Tetanus toxin frag
41	114	57.0	453	22	AAB31427	Amino acid sequenc
42	114	57.0	463	20	AAY00921	Tetanus toxin frag
43	114	57.0	573	8	AA170345	Portion of B fragm
44	114	57.0	605	22	AAE07897	Modified clostridi
45	114	57.0	618	19	AAW48909	SOD-1/TTC hybrid p

ALIGNMENTS

RESULT 1
AAB20150
ID AAB20150 standard; Protein; 109 AA.
XX
AC AAB20150;
XX
30-APR-2001 (first entry)
XX
Growth differentiation factor 8 AutoVac construct GDF-8 P30-3A.
XX
Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
Chimeric - Homo sapiens.
OS Chimeric - Clostridium tetani.
OS Synthetic.
XX
Key Location/Qualifiers
FT Region 1..78
FT /note= "identical to residues 267-345 of human GDF-8"
FT
FT Region 79..99
FT /note= "tetanus toxoid P2 epitope"
FT Region 100..109
FT /note= "identical to residues 366-375 of human GDF-8"
FT
FT Misc-difference 73
FT /note= "Cys-73 may be substituted by Ser to avoid disulfide bond formation"
FT
FT Misc-difference 90..91
FT /note= "optionally replaced by Glu-Gly"
XX
WO200105820-A2.

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC		EMBL; J03203; AAA29553.1; -.	
DR	DR	PIR; A28615; A28615.	
DR	DR	HSSP; P04002; LWFA.	
DR	DR	InterPro; IPR000884; TSP1.	
DR	DR	Pfam; PF00090; tsp_1; 1.	
DR	DR	SMART; SM00209; TSP1; 1.	
DR	DR	PROSITE; PS0092; TSP1; 1.	
KW	Malaria;	Repeat; Sporozoite.	
FT	NON_TER	1	
FT	DOMAIN	16	278
SQ	SEQUENCE	393 AA;	TANDEM REPEATS OF G-N-A-A. B32944419BC600AA CRC64;

Query Match	22.3%	Score 59;	DB 1;	Length 393;
Best Local Similarity	50.0%;	Pred. No. 7.2;		
Matches 11; Conservative	6;	Mismatches	5;	Indels 0;
Gaps	0;			

Qy 18 EKKIAKMEKASSVFNVNSSG 39
 | : | : | : | : | : |
Dd 361 ETEICSLDKCSSIFNVVNSLG 382

Search completed: October 10, 2002, 16:06:39
Job time : 7.64744 secs


```
RA Rodriguez R.E., Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=72094314; PubMed=4550508;
RA Burqus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vaie W., Guillemin R.;
RT "Primary structure of the ovine hypothalamic luteinizing hormone-
RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-bdman degradation).";
RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U02517; AAA03433.1; -.
DR PIR: A93780; RSHSG.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
CC
CC EMBL: U02517; AAA03433.1; -.
DR PIR: A93780; RSHSG.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >61 PROGONADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;

Query Match 23.28; Score 61.5; DB 1; Length 61;
Best Local Similarity 42.98; Pred. No. 0.45;
Matches 15; Conservative 3; Mismatches 2; Indels 15; Gaps 2;

Qy 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVFNVNS 36
| | | | | | | | | | | | | | | | | | | |
Db 2 HWSYGLRPGG-----KRNK-----NVIDS 21

RESULT 12
GON3_SPAAU STANDARD; PRT; 90 AA.
AC P51923;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberein III precursor (Gonadotropin-releasing hormone III)
DE (GNRH-III) (LH-RH III) (Luliberin III).
GN GNRH3.
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99061849; PubMed=9843645;
RA Holland M.C.H., Gothliff Y., Meiri I., King J.A., Okuzawa K.,

RA Elizur A., Zohar Y.;
RT "Levels of the native forms of GnRH in the pituitary of the gilthead
RT seabream, Sparus aurata, at several characteristic stages of the
RT gonadal cycle.";
RL Gen. Comp. Endocrinol. 112:394-405(1998).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U30311; AAA98845.1; -.
DR InterPro: IPR002012; GNRH.
DR Pfam: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 90 PROGONADOLIBERIN III.
FT PEPTIDE 24 33 GONADOLIBERIN III.
FT PEPTIDE 37 82 GNRH-ASSOCIATED PEPTIDE III (POTENTIAL).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT SEQUENCE 90 AA; 10030 MW; FB4E47EB868C2FBD CRC64;

Query Match 23.28; Score 61.5; DB 1; Length 90;
Best Local Similarity 30.48; Pred. No. 0.69;
Matches 17; Conservative 6; Mismatches 16; Indels 17; Gaps 2;

Qy 2 HWSYGLRPG-----SSGP-----SLDEKKIAKMEKASSVFNVNSSCP 40
| | | | | | | | | | | | | | | | | | | |
Db 25 HWSYGLPGRKSRVGELEATIRMGTVGVSLPEEASQAQTOERLPYNNIKDDSSP 80
| | | | | | | | | | | | | | | | | | | |

RESULT 13
GON3_DICLA STANDARD; PRT; 90 AA.
AC Q91A09;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberein III precursor (Gonadotropin-releasing hormone III)
DE (GNRH-III) (LH-RH III) (Luliberin III).
GN GNRH3.
OS Dicotylarchus labrax (European sea bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Moronidae; Dicotylarchus.
OX NCBI_TaxID=13489;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanuy S.,
RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;
RT "Differential expression of three different prepro-GNRH
RT (Gonadotropin-releasing hormone) messengers in the brain of the
RT European sea bass (Dicotylarchus labrax).";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```

Oy 18 EKKIATMEKASSVFNVNSSG 39
Db 335 DTECKMDKCSIFNIVNSLG 356

RESULT 9
GON3_SALTR STANDARD; PRT; 82 AA.
AC P45653;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadoliberin III precursor (Gonadotropin-releasing hormone III)
DE (GNRH-III) (LHRH III) (Luliberin III).
GN GNRH3.
OS Salmo trutta (Brown trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8032;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Klungland H., Anderson O., Alestrom P.;
RT "the salmon gonadotropin-releasing hormone encoding gene in
RL Mol. Mar. Biol. Biotechnol. 1:420-425(1992).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; X79713; CAA56152.1; -.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 82 PROGNADOLIBERIN III.
FT PEPTIDE 24 83 GONADOLIBERIN III.
FT PEPTIDE 37 82 GNRH-ASSOCIATED PEPTIDE III (POTENTIAL).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 82 AA; 9191 MW; 8053E9534A765408 CRC64;

Query Match 23.6%; Score 62.5; DB 1; Length 82;
Best Local Similarity 32.7%; Pred. No. 0.47;
Matches 17; Conservative 6; Mismatches 20; Indels 9; Gaps 2;

Oy 2 HWSYGLRPGSGPSLDEKKIAKME---KASSVFNVNSSGPSLHWSYGLRP 50
Db 25 HWSYGLWLPGG-----KRSVGELEATIKMMDTGCVVALPEETSAHFSERLRP 70

RESULT 10
GON1_MESAU STANDARD; PRT; 63 AA.
AC O09163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I]
DE (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN=WESTERN RANGE; TISSUE=Hypothalamus;

```

```

GN GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; U91938; AAB51302.1; -.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >63 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY
FT ACT_SITE 3 3 SIMILARITY).
FT MOD_RES 1 1 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT MOD_RES 10 10 ACTIVITY (BY SIMILARITY).
FT MOD_RES 63 63 PYRROLIDONE CARBOXYLIC ACID (BY
FT NON_TER 63 63 SIMILARITY).
FT SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;

Query Match 23.4%; Score 62; DB 1; Length 63;
Best Local Similarity 33.3%; Pred. No. 0.4;
Matches 18; Conservative 5; Mismatches 9; Indels 22; Gaps 4;

Oy 2 HWSYGLRPGSS-----GPSLDEKKIAKMEKASSVFNVNSSGSP-----SLHW 44
Db 2 HWSYGLRPGGKRNARLGDSPQE-----MDKE-----VDQLAEPQHLECTVHW 44

RESULT 11
GON1_SHEEP STANDARD; PRT; 61 AA.
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN=WESTERN RANGE; TISSUE=Hypothalamus;

```


[illegible]

```

01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5846;
[1]
SEQUENCE FROM N.A.
MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S.; Enea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from
Thailand.";
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE
MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE
VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR
ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES
WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; M19752; AAA29555.1; -.
DR PIR; A54533;
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSRZTOITE.
DR SMART; SM00209; TSPL; 1.
DR Malaria; Sporozoite; Repeat; Signal.
KW SIGNAL 1 16 PROBABLE.
FT CHAIN 17 424 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 300 45 X 4 AA TANDEM REPEATS OF N-A-N-P.
SQ SEQUENCE 424 AA; 43610 MW; 710AB14238786CD9 CRC64;

Query Match 37.4%; Score 99; DB 1; Length 424;
Best Local Similarity 56.1%; Pred. No. 9.3e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGPSLD-----EKKIAKMEKASSVFNVNVSSEG 39
      :|||: | ||| ||| ||||||||| |
Db 373 IKPGANKPKDELVDYEIEKKICKMEKCSSFVNFVNNSIG 413

RESULT 5
CSP_PLAFLW STANDARD; PRT; 442 AA.
ID AC P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5848;
[1]
SEQUENCE FROM N.A.
MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J.; Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium
falciparum.";
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT
SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 6.64744 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-20

Perfect score: 265

Sequence: 1 XHWSYGLRPSSGSLDEKK.....NVNSSSGPSLHWSYGLRPX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	105	39.6	388	1 CSP_PLARE	P26694 plasmodium
2	99	37.4	397	1 CSP_PLAFO	P19597 plasmodium
3	99	37.4	412	1 CSP_PLAFA	P02893 plasmodium
4	99	37.4	424	1 CSP_PLAFT	P13814 plasmodium
5	98	37.0	442	1 CSP_PLAFW	P08307 plasmodium
6	66	24.9	339	1 CSP_PLABE	P06915 plasmodium
7	66	24.9	347	1 CSP_PLABA	P23093 plasmodium
8	63	23.8	367	1 CSP_PLAYO	P06914 plasmodium
9	62.5	23.6	82	1 GON3_SALTR	P45653 salmo trutt
10	62	23.4	63	1 GON1_MESAU	O09163 mesocricetu
11	61.5	23.2	61	1 GON1_SHEEP	Q26588 ovis aries
12	61.5	23.2	90	1 GON3_SPAAU	P51923 sparus aura
13	60.5	22.8	90	1 GON3_DICLA	Q91a09 dicentrarch
14	60	22.6	94	1 GON1_HAPBU	P51918 haplochromi
15	59	22.3	393	1 CSP_PLABR	P14593 plasmodium
16	59	22.3	429	1 CSP_PLAMA	P13815 plasmodium
17	58.5	22.1	74	1 GON3_ONCTS	Q20209 oncorhynch
18	58.5	22.1	82	1 GON3_SALSA	P35629 salmo salar
19	58	21.9	67	1 GON1_MACMU	P55247 macaca mula
20	58	21.9	89	1 GON1_XENLA	P45656 xenopus lae
21	58	21.9	90	1 GON1_MOUSE	P13562 mus musculu
22	58	21.9	91	1 GON1_PIG	P49921 sus scrofa
23	58	21.9	92	1 GON1_HUMAN	P01148 homo sapien
24	58	21.9	92	1 GON1_RAT	P07490 rattus norv
25	58	21.9	92	1 GON1_TURGB	Q95335 tupaia glis
26	58	21.9	464	1 TRME_BORBU	P53364 borrelia bu
27	57.5	21.7	74	1 GON3_ONCMY	P55246 oncorhynch
28	57.5	21.7	90	1 GON3_HAPBU	P45652 haplochromi
29	57.5	21.7	3133	1 HMCT_BORMO	P98092 bombyx mori
30	57	21.5	309	1 FDHE_ECOLI	P13024 escherichia
31	56.5	21.3	184	1 DEF2_BAGST	O31410 bacillus st
32	56.5	21.3	404	1 IIBC_HUMAN	P29466 homo sapien
33	55.5	20.9	89	1 GON3_PORNO	P51922 porichthys

ALIGNMENTS

RESULT 1

CSP_PLARE STANDARD; PRT; 388 AA.

AC P26694;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
RX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91201303; PubMed=2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria parasite evolutionarily related to the human malaria parasite Plasmodium falciparum.";
RL J. Biol. Chem. 266:6686-6689(1991).
CC -!- FUNCTION: THE CIRCUMSPOROZOITE PROTEIN IS THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE (THE INFECTIVE STAGE OF THE MALARIA PARASITE THAT IS TRANSMITTED FROM THE MOSQUITO TO THE VERTEBRATE HOST).
CC -!- MISCELLANEOUS: THE C-TERMINAL REGION IS PROBABLY USED FOR ANCHORING THE PROTEIN TO THE CELL MEMBRANE. THE REPEAT SEQUENCES WOULD BE THE SURFACE ANTIGEN OF THE ORGANISM.
CC -!- SIMILARITY: CONTAINS 1 TSP TYPE-1 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60972; AAA29561.1; -;
DR PIR; A39756; A39756;
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF000090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
KW Malaria; Sporozoite; repeat; Signal.
FT SIGNAL 1
FT CHAIN 17 388
FT DOMAIN 120 267
SQ SEQUENCE 388 AA; 42245 MW; C031EEFBE2E35604 CRC64;
CIRCUMSPOROZOITE PROTEIN.
62 X 4 AA TANDEM REPEATS OF N-A-N-P.

Query Match 39.6%; Score 105; DB 1; Length 388;
Best Local Similarity 58.5%; Pred. No. 1.5e-05;
Matches 24; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGSPSLD-----EKKIARMEKASSVFNVNSSSG 39

Db 337 IKPSAGKPRDQLDYENDLEKKICKMEKSSVFNVNSSIG 377

Db 397 ETICSLDKCSSIFNVVNSLG 418

Search completed: October 10, 2002, 16:12:17
Job time : 12.9551 secs

	Matches	19;	Conservative	6;	Mismatches	22;	Indels	25;	Gaps	2;
Qy	4	SYGLRPGSSGPSLDEKKIAKMEKASSYFN-----VVN-----SSS	38							
		I : I :								
Db	1448	SLSIIMGSSGGPPDYDRAHVTGACSSSSSTKTGYFPAILNPSPATERSHYTMEFGYSN	1507							
		I : I :								
Qy	39	GPSLHWISYGLRP	50							
		I I I I								
Db	1508	SPSTHRYSYRYP	1519							
		I I I I								
RESULT 14										
JE0273										
		low density lipoprotein receptor-related protein 6 - mouse								
C:	Species:	Mus musculus (house mouse)								
C:	Date:	05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 18-Aug-2000								
C:	Accession:	JE0273								
R:	Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metz									
A:	Title:	Isolation and characterization of LRP6, a novel member of the low density li								
A:	Reference number:	JE0272; MUID:98369644								
A:	Accession:	JE0273								
A:	Molecule type:	mRNA								
A:	Residues:	1-1613 <BRO>								
A:	Cross-references:	GB:AF074265; NID:g3462528; PIDN:AAC33007.1; PID:g3462529								
C:	Genetics:									
A:	Gene:	Lrp6								
A:	Map position:	6								
C:	Superfamily:	unassigned EGF-related proteins; EGF homology; LDL receptor ligand-bin								
F:	286-323/Domain:	EGF homology <EGF1>								
F:	592-627/Domain:	EGF homology <EGF>								
F:	1207-1243/Domain:	EGF homology <EGF2>								
F:	1249-1285/Domain:	LDL receptor ligand-binding repeat homology <LDL1>								
F:	1288-1322/Domain:	LDL receptor ligand-binding repeat homology <LDL>								
F:	1326-1360/Domain:	LDL receptor ligand-binding repeat homology <LDL3>								
Query Match		22.5%; Score 59.5; DB 2; Length 1613;								
Best Local Similarity		26.4%; Pred. No. 1e+02;								
Matches	19;	Conservative	6;	Mismatches	22;	Indels	25;	Gaps	2;	
Qy	4	SYGLRPGSSGPSLDEKKIAKMEKASSYFN-----VVN-----SSS	38							
		I : I :								
Db	1448	SLSIIMGSSGGPPDYDRAHVTGACSSSSSTKTGYFPAILNPSPATERSHYTMEFGYSN	1507							
		I : I :								
Qy	39	GPSLHWISYGLRP	50							
		I I I I								
Db	1508	SPSTHRYSYRYP	1519							
		I I I I								
RESULT 15										
A54504										
		circumsporozoite protein - Plasmodium malariae (strain Uganda 1-CDC)								
C:	Species:	Plasmodium malariae								
C:	Date:	06-Oct-1994 #sequence_revision 04-Nov-1994 #text_change 20-Aug-1999								
C:	Accession:	A54504								
R:	Lal, A.A.; de la Cruz, V.F.; Campbell, G.H.; Procell, P.M.; Collins, W.E.; McCutcha									
Mol.	Biochem. Parasitol.	30, 291-294, 1988								
A:	Title:	Structure of the circumsporozoite gene of Plasmodium malariae.								
A:	Reference number:	A54504; MUID:89040027								
A:	Accession:	A54504								
A:	Molecule type:	DNA								
A:	Residues:	1-429 <LAL>								
A:	Cross-references:	GB:J03992; NID:g160220; PIDN:AAA29557.1; PID:g160221								
C:	Superfamily:	circumsporozoite protein; thrombospondin type 1 repeat homology								
C:	Keywords:	tandem repeat								
F:	354-407/Domain:	thrombospondin type 1 repeat hom								

RESULT 9

OZZOMY

circumsporozoite protein precursor - Plasmodium yoelii

N:Alternate names: sporozoite surface antigen

C:Species: Plasmodium yoelii

C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999

C:Accession: A26271

R:Lal, A.A.; de la Cruz, V.F.; Welsh, J.A.; Charoenvit, Y.; Maloy, W.L.; McCutchan, T. J. Biol. Chem. 262, 2937-2940, 1987

A:Title: Structure of the gene encoding the circumsporozoite protein of Plasmodium yoelii

A:Reference number: A26271; MUID:87137555

A:Accession: A26271

A:Molecule type: DNA

A:Residues: 1-367 <IAL>

A:Cross-references: GB:J02695; NID:g160222; PIDN:AAA29558.1; PID:g160223

C:Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.

C:Comment: There are 15 copies of a 6-residue repeat and 8 copies of a 4-residue repeat

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

C:Keywords: sporozoite; surface antigen; tandem repeat

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-367/Product: circumsporozoite protein #status predicted <MAT>

F:139-228/Region: 6-residue repeats (Q-G-P-G-A-P)

F:229-260/Region: 4-residue repeats (Q-Q-P-P)

F:293-345/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 23.8%; Score 63; DB 1; Length 367;

Best Local Similarity 50.0%; Pred. No. 7.3;

Matches 11; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 18 EKKIAKMEKASSVFNVNSSG 39

: : | | | | | | | | | | | |

DB 335 DTEICKMDKCSIFNIVSNL 356

RESULT 10

I51365

gonadotropin-releasing hormone - brown trout

C:Species: Salmo trutta (brown trout)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: I51365

R:Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P. Mol. Cell. Endocrinol. 84, 167-174, 1992

A:Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.

A:Reference number: I51040; MUID:92267241

A:Accession: I51365

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-82 <KLU>

A:Cross-references: EMBL:X79713; NID:g499343; PIDN:CAA56152.1; PID:g499344

C:Genetics:

A:Gene: GnRH

A:Introns: 46/3; 73/3

Query Match 23.6%; Score 62.5; DB 2; Length 82;

Best Local Similarity 32.7%; Pred. No. 1.6;

Matches 17; Conservative 6; Mismatches 20; Indels 9; Gaps 2;

QY 2 HWSYGLRPGSSGLDKKIAKME--KASSVFNVNSSSGPSLHWSYGLRP 50

||||| | | | | | | | | | | | | | | |

DB 25 HWSYGLPVG-----KRSVGELEATIKMMDTGCVLPETSAHFSERL 70

RESULT 11

I50739

gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)

C:Species: Haplochromis burtoni

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: I50739

R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D. Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995

RESULT 5

A54529
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain Wellcome)
C:Species: Plasmodium falciparum
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54529
R:Lockyer, M.J.; Schwarz, R.T.
Mol. Biochem. Parasitol. 22, 101-108, 1987
A>Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum
A:Reference number: A54529; MUID:87115616
A:Accession: A54529
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-442 <LOC>
A:Cross-references: GB:M15505; NID:gl60214; PIDN:AAA29554.1; PID:gl60215
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
F:366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 37.0%; Score 98; DB 2: Length 442;
Best Local Similarity 56.1%; Pred.No. 0.00055;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVNVNVSSTG 39
:||||| | |||| |||| |||| |||| |||| |||| ||||
DB 391 IKPGSADKPQLDYNDIEKKIKMEKCSVNVNVSSTG 431

RESULT 6

OZZQMB
Circumsporozoite protein precursor - Plasmodium berghei (strain NK65)
N:Alternate names: sporozoite surface antigen
C:Species: Plasmodium berghei
C>Date: 30-Sep-1987 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999
C:Accession: A44948; A25083; S13446
R:Lanar, D.E.
Mol. Biochem. Parasitol. 39, 151-154, 1990
A>Title: Sequence of the circumsporozoite gene of Plasmodium berghei ANKA clone and N
A:Reference number: A44948; MUID:90158693
A:Accession: A44948
A:Molecule type: DNA
A:Residues: 1-332 <LAN>
A:Cross-references: GB:M28887
R:Eichinger, D.J.; Arnot, D.E.; Tam, J.P.; Nussenzweig, V.; Enea, V.
Mol. Cell. Biol. 6, 3965-3972, 1986
A>Title: Circumsporozoite protein of Plasmodium berghei: gene cloning and identification
A:Reference number: A25083; MUID:87089740
A:Accession: A25083
A:Molecule type: DNA
A:Residues: 1-26,'I',28-68,'PMLRR','75-126','P',128-134,'PPNPANDP',135-332 <EIC>
A:Cross-references: GB:M14135; NID:gl60245; PIDN:AAA29577.1; PID:gl60246
R:Weber, J.L.; Egan, J.E.; Lyon, J.A.; Wirtz, R.A.; Charoenvit, Y.; Maloy, W.L.; Hock
Exp. Parasitol. 63, 295-300, 1987
A>Title: Plasmodium berghei: cloning of the circumsporozoite protein gene.
A:Reference number: S13446; MUID:87218962
A:Accession: S13446
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 61-122,'A',124-332 <WEB>
A:Cross-references: GB:M25445; NID:gl60177; PIDN:AAA29531.1; PID:gl60178
C:Comment: There are three distinct regions in the mature circumsporozoite protein, t
obic membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: sporozoite; surface antigen; tandem repeat
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-32/Product: circumsporozoite protein #status predicted <MAT>
F:94-189/Region: 8-residue repeats
F:199-230/Region: 2-residue repeats
F:258-310/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 24.9%; Score 66; DB 1: Length 332;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 11.8782 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834a-20

Perfect score: 265

Sequence: 1 XHSYGLRPGSGPSLDEKK.....NVNSSGSLHWSYGLRPX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	105	39.6	388	2 A39756	circumsporozoite p
2	99	37.4	405	2 S05428	circumsporozoite p
3	99	37.4	412	1 OZZQAF	circumsporozoite p
4	99	37.4	424	2 A54533	circumsporozoite p
5	98	37.0	442	2 A54529	circumsporozoite p
6	66	24.9	332	1 OZZQMB	circumsporozoite p
7	66	24.9	348	1 OZZQBK	circumsporozoite p
8	63	23.8	264	2 A44969	circumsporozoite p
9	63	23.8	367	1 OZZQMY	circumsporozoite p
10	62.5	23.6	82	2 I51365	gonadotropin-relea
11	60	22.6	98	2 I50739	gonadotropin-relea
12	59.5	22.5	315	2 F98295	hypothetical prote
13	59.5	22.5	1613	2 JE0272	low density lipopr
14	59.5	22.5	1613	2 JE0273	low density lipopr
15	59	22.3	429	2 A54504	circumsporozoite p
16	59	22.3	485	2 A60610	circumsporozoite p
17	58.5	22.1	74	2 I51092	gonadotropin relea
18	58.5	22.1	82	2 I51355	gonadotropin relea
19	58.5	22.1	82	2 I51331	gonadotropin relea
20	58.5	22.1	624	2 T38006	probable lysophosp
21	58	21.9	10	1 RHPQG	gonadoliberin - pi
22	58	21.9	10	1 RSHSG	gonadoliberin - sh
23	58	21.9	67	2 I78541	gonadoliberin prec
24	58	21.9	89	2 I51423	gonadoliberin prec
25	58	21.9	90	1 RHMSG	gonadoliberin prec
26	58	21.9	92	1 RHHUG	gonadoliberin prec
27	58	21.9	92	1 RHRTG	gonadoliberin prec
28	58	21.9	345	2 T29416	hypothetical prote
29	58	21.9	464	2 C70122	thiophene and fura

30 57.5 21.7 90 2 A23735 gonadoliberin prec
31 57.5 21.7 3133 2 S52093 hemocytin - silkwo
32 57 21.5 309 2 S40835 fdHE protein - Esc
33 57 21.5 309 2 A91231 affects formate de
34 57 21.5 309 2 H86077 affects formate de
35 57 21.5 382 2 A83215 conserved hypothet
36 56.5 21.3 263 2 C56084 interleukin-1beta
37 56.5 21.3 287 2 T19917 hypothetical prote
38 56.5 21.3 311 2 B56084 interleukin-1beta
39 56.5 21.3 383 2 A56084 interleukin-1beta
40 56.5 21.3 404 2 A42677 interleukin-1 beta
41 56.5 21.3 561 2 T35845 hypothetical prote
42 55.5 20.9 90 2 JC7395 salmon-type gonado
43 55.5 20.9 580 2 T32851 hypothetical prote
44 55.5 20.9 673 2 T50281 probable lysophosp
45 55 20.8 249 2 A41497 36K antigen pra -

ALIGNMENTS

RESULT 1

A39756 circumsporozoite protein - Plasmodium reichenowi

C;Species: Plasmodium reichenowi

C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999

C;Accession: A39756

R;Lal, A.A.; Goldman, I.F.

J. Biol. Chem. 266, 6686-6689, 1991

A;Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malar

A;Reference number: A39756; MUID:91201303

A;Accession: A39756

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-388 <LAL>

A;Cross-references: GB:M60972; NID:gl60228; PIDN:AAA29561.1; PID:gl60229

C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F;312-366/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 39.6%; Score 105; DB 2; Length 388;

Best Local Similarity 58.5%; Pred. No. 6.9e-05;

Matches 24; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPSSGPSLD-----EKKIAMEKASSVFNVNSSSG 39

Db 337 IKPGSAGKPKDQLDYENDLEKKICKMEKCSSVFNVNSSIG 377

RESULT 2

S05428

circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)

C;Species: Plasmodium falciparum

C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000

C;Accession: S05428; A45527; I60657

R;Campbell, J.R.

Nucleic Acids Res. 17, 5854, 1989

A;Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate

A;Reference number: S05428; MUID:89345189

A;Accession: S05428

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-405 <CAM>

A;Cross-references: EMBL:X15363

R;Caspers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.

Mol. Biochem. Parasitol. 35, 185-190, 1989

A;Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate

A;Reference number: A45527; MUID:89364998

A;Accession: A45527

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-405 <CAS>

A;Cross-references: GB:M22982; GB:J04650; NID:gl60168; PIDN:AAA29527.1; PID:gl60169

R;Lockyer, M.J.; Marsh, K.; Newbold, C.I.

Db 495 HWSYGLRPGSGSQDWSY-----GLR--PGSSQHWSYGLRP 528

RESULT 15

US-08-387-156-8
; Sequence 8, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-8

Query Match 43.48; Score 106; DB 1; Length 977;
Best Local Similarity 50.08; Pred. No. 1.8e-05;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

Qy 2 HWSYGLRPGSGSPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
|||||
Db 928 HWSYGLRPGSGSQDWSY-----GLR--PGSSQHWSYGLRP 961

Search completed: October 10, 2002, 16:14:06
Job time : 9.73077 secs

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-878-748-10

Query Match 43.4%; Score 106; DB 2; Length 544;
Best Local Similarity 50.0%; Pred No. 9,2e-06;
Matches 22; Conservative 1; Mismatches 11; Indels

QY 2 HWSYGLRPGSGPSLQYIKANSKFGITELSSGPSLHWSYGLRP 45
|||||
Db 495 HWSYGLRPGSGQWSY-----GLR--PGSSQHSYGLRP 528

RESULT 14
US-09-124-491-10
; Sequence 10, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-124-491-10

Query Match 43.4%; Score 106; DB 3; Length 544;
Best Local Similarity 50.0%; Pred No. 9,2e-06;
Matches 22; Conservative 1; Mismatches 11; Indels

QY 2 HWSYGLRPGSGPSLQYIKANSKFGITELSSGPSLHWSYGLRP 45
|||||

```

```
; MOLECULE TYPE: protein
US-08-878-748-4
Query Match 43.4%; Score 106; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 5.6e-07;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGSLQYIKANSKFITELSSGSPSLHWSYGLRP 45
|||||
Db 2 HWSYGLRPGSGSQDWSY-----GLR--PGSSQHWSYGLRP 35

RESULT 10
US-09-124-491-4
; Sequence 4, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; APPLICATION NUMBER: US/09124,491
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-124-491-4
Query Match 43.4%; Score 106; DB 3; Length 49;
Best Local Similarity 50.0%; Pred. No. 5.6e-07;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGSLQYIKANSKFITELSSGSPSLHWSYGLRP 45
|||||
Db 2 HWSYGLRPGSGSQDWSY-----GLR--PGSSQHWSYGLRP 35

RESULT 11
US-09-124-491-4
; Sequence 4, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; APPLICATION NUMBER: US/09124,491
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-124-491-4
Query Match 43.4%; Score 106; DB 3; Length 49;
Best Local Similarity 50.0%; Pred. No. 5.6e-07;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGSLQYIKANSKFITELSSGSPSLHWSYGLRP 45
|||||
Db 2 HWSYGLRPGSGSQDWSY-----GLR--PGSSQHWSYGLRP 35

RESULT 12
US-08-694-865-10
; Sequence 10, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; APPLICATION NUMBER: US/08387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-10
Query Match 43.4%; Score 106; DB 1; Length 544;
Best Local Similarity 50.0%; Pred. No. 9.2e-06;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGSLQYIKANSKFITELSSGSPSLHWSYGLRP 45
|||||
Db 495 HWSYGLRPGSGSQDWSY-----GLR--PGSSQHWSYGLRP 528
```

```

; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS, P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-694-865-4

Query Match 43.4%; Score 105; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 5.6e-07;
Matches 22; Conservative 1; Mismatches 11; Indels

QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGPSLHWSYGLRP 45
||||| | | | | | | | | | | | | | | | | | | | | |
Db 2 HWSYGLRPGSGSDWSY-----GLR--PGSSQHSYGLRP 35

RESULT 9
US-08-787-748-4
; Sequence 4, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

Query Match 44.5%; Score 108.5; DB 2; Length 699;

RESULT 7
US-08-387-156-4
; Sequence 4, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:

2

PD 19-FEB-1998.
 XX
 PF 08-AUG-1997; 97WO-CA00559.
 XX
 PR 09-AUG-1996; 96US-0694865.
 XX
 PA (UUSA-) UNIV SASKATCHEWAN.
 XX
 PI Manns JG, Potter AA;
 XX
 PP WPI; 1998-159540/14.
 XX
 DR N-PSDB; AAV61531.
 XX
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -
 PT useful for, e.g. preparation of vaccines for reduction of incidence
 PT of mammary tumours in mammals
 XX
 PS Disclosure; Figure 5.1-8; 118pp; English.
 XX
 CC The present sequence represents the LKT-GnRH chimeric protein from
 CC pCB113. This plasmid contains the LKT 352 polypeptide (AAW79568) fused
 CC to four copies of the GnRH peptide. This chimera lacks cytotoxic
 CC activity which enables there to be an increase in antigen presentation
 CC and thus an optimal immune response. The removal of this region also
 CC enables the truncated LKT to be expressed at much higher levels and
 CC allows the amount of antigen administered to be reduced. This chimeric
 CC protein comprises a leukotoxin polypeptide, several multimers, and a GnRH
 CC sequence. The chimeric protein can be used as a vaccine to help reduce
 CC the incidence of mammary tumours in a mammalian individual.
 XX
 SQ Sequence 977 AA;
 Query Match 43.4%; Score 106; DB 19; Length 977;
 Best Local Similarity 50.0%; Pred. No. 5.7e-05;
 Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;
 QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSLHWSYGLRP 45
 |||||
 Db 928 HWSYGLRPGSGQWSY-----GLR--PGSSQHWSYGLRP 961
 |||||
 RESULT 15
 AAB20865
 ID AAB20865 standard; peptide; 42 AA.
 XX
 AC AAB20865;
 XX
 DT 03-JAN-2001 (first entry)
 XX
 DE GnRH tandem dimer peptide sequence SEQ ID NO:3.
 XX
 KW Gonadotropin releasing hormone; GnRH; immunogen; Protein D; carrier;
 KW prostate cancer; Haemophilus influenzae; vaccine; infectious disease;
 KW malaria; cytostatic; antiallergic; neuroprotective;
 KW protozoacide; Alzheimer's disease; allergy.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 FH Modified-site 42
 FT /note= "amidated"
 XX
 PN WO200050077-A1.
 XX
 PD 31-AUG-2000.
 XX
 PF 22-FEB-2000; 2000WO-EP01457.
 XX
 PR 25-FEB-1999; 99GB-0004405.
 PR 25-FEB-1999; 99GB-0004408.
 PR 25-FEB-1999; 99GB-0004412.
 PR 13-AUG-1999; 99GB-0019260.
 XX

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Coste M, Lobet Y, Van-Mechelen MP, Verliest C;
 XX
 DR WPI; 2000-572040/53.
 XX
 XX Immunogens and vaccine comprising the immunogen useful for preventing
 PT and treating infectious diseases e.g. malaria and chronic disease e.g.
 PT cancer, comprises peptide and carrier from protein D of influenzae -
 XX
 PS Disclosure; Page 7; 53pp; English.
 XX
 CC The present invention describes an immunogen (I) comprising a peptide
 CC (Ia) and a carrier (Ib) derived from protein D of Haemophilus influenzae
 CC or its fragment. Also described are: (1) a vaccine comprising (I), and
 CC an excipient; (2) preparation of (I), comprising conjugating a peptide
 CC to protein D or its fragment; and (3) preparation of a vaccine of (1),
 CC comprising formulating (I) with an excipient. (I) has cytostatic,
 CC antiallergic, nootropic, neuroprotective and protozoacide activities.
 CC (1) and the vaccine are useful for the manufacture of a medicament for
 CC preventing and treating infectious diseases such as malaria or chronic
 CC disease such as cancer, Alzheimer's disease or allergy in a patient.
 CC Unlike prior art immunogens, (I) induces high levels of antipeptide
 CC immune responses while inducing a moderate humoral response against the
 CC carrier. The present sequence represents an example of an immunogen from
 CC the present invention which contains gonadotropin releasing hormone
 CC (GnRH) tandem dimers.
 XX
 SQ Sequence 42 AA;
 Query Match 38.5%; Score 94; DB 21; Length 42;
 Best Local Similarity 45.7%; Pred. No. 4.3e-05;
 Matches 21; Conservative 0; Mismatches 7; Indels 18; Gaps 2;
 QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSG--PSLHWSYGLRP 45
 |||||
 Db 12 HWSYGLRPGS-----CEHWSYGLRPGQHWSYGLRP 41
 |||||
 Search completed: October 10, 2002, 16:05:16
 Job time : 24.1795 secs

```

RESULT 12
AAW79570
ID AAW79570 standard; Protein; 544 AA.
XX AC AAW79570;
XX XX
XX DT 24-DEC-1998 (first entry)
XX DE LKT-GnRH chimeric protein.
XX XX
XX KW Chimera; pCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
XX KW cytotoxic activity; antigen presentation; immune response; vaccine;
XX KW tumour.
XX OS Synthetic.
XX XX
XX PN WO9806848-A1.
XX XX
XX PD 19-FEB-1998.
XX XX
XX PF 08-AUG-1997; 97WO-CA00559.
XX XX
XX PR 09-AUG-1996; 96US-0694865.
XX XX
XX PA (UYSA-) UNIV SASKATCHEWAN.
XX XX
XX PI Manns JG, Potter AA;
XX XX
XX DR WPI; 1998-159540/14.
XX DR N-PSDB; AAV61532.
XX XX
XX PT Chimeric protein of leukotoxin and gonadotropin releasing hormone -
XX PT useful for, e.g. preparation of vaccines for reduction of incidence
XX PT of mammary tumours in mammals
XX PS Disclosure; Figure 7.1-5; 118pp; English.
XX XX
XX CC The present sequence represents the LKT-GnRH chimeric protein from
XX CC pCB111. This plasmid contains the LKT 111 polypeptide fused to
XX CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity
XX CC which enables there to be an increase in antigen presentation and thus an
XX CC optimal immune response. The removal of this region also enables the
XX CC truncated LKT to be expressed at much higher levels and allows the amount
XX CC of antigen administered to be reduced. This chimeric protein comprises a
XX CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
XX CC chimeric protein can be used as a vaccine to help reduce the incidence of
XX CC mammary tumours in a mammalian individual.
XX SQ Sequence 544 AA;

Query Match 43.4%; Score 106; DB 19; Length 544;
Best Local Similarity 50.0%; Pred. No. 2.8e-05;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGPSLOYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 495 HWSYGLRPGSGQDWSY-----GLR--PGGSSQHWSYGLRP 528

RESULT 13
AAW03942
ID AAW03942 standard; Protein; 977 AA.
XX AC AAW03942;
XX XX
XX DT 20-NOV-1996 (first entry)
XX DE LKT-GnRH protein fusion from pCB113.
XX XX
XX KW Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
XX KW fusion protein; immunogen; vaccine; fertility control;

contraceptive; sterilisation;
Chimeric Pasteurella haemolytica A1 strain B122;
Chimeric synthetic.
Key Location/Qualifiers
Domain 1..929
Domain /label= LKT
Domain 927..977
Domain /label= GnRH_repeat_domain
WO9624675-A1.
15-AUG-1996.
24-JAN-1996; 96WO-CA00049.
10-FEB-1995; 95US-0387156.
(UYSA-) UNIV SASKATCHEWAN.
Manns JG, Potter AA;
WPI; 1996-384447/38.
N-PSDB; AAT37176.
Gonadotropin-releasing hormone multimer fusion proteins - with
leukotoxin polypeptide for increased immunogenicity, useful in
antifertility vaccine prodn.
Claim 7; Fig 5A-5H; 87pp; English.
A chimeric protein (AAW03942) is composed of a fusion between
a truncated leukotoxin (LKT-352) from Pasteurella haemolytica (see
also AAW03945) and a 4-copy gonadoliberein-releasing hormone (GnRH)
chimeric gene (AAT37176) produced by ligating a synthetic sequence
for the 4-copy GnRH into vector pAA352 (ATCC 68283), which carries
the LKT-352 gene. Recombinant plasmid pCB113 (LKT 352:4 copy
GnRH, ATCC 69749) was obtd. Escherichia coli transformants
produced the chimeric protein, which is useful as a vaccine for
fertility control, esp. immunological sterilisation of
domestic or farm animals.
SQ Sequence 977 AA;

Query Match 43.4%; Score 106; DB 17; Length 977;
Best Local Similarity 50.0%; Pred. No. 5.7e-05;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGPSLOYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 928 HWSYGLRPGSGQDWSY-----GLR--PGGSSQHWSYGLRP 961

RESULT 14
AAW79569
ID AAW79569 standard; Protein; 977 AA.
XX AC AAW79569;
XX XX
XX DT 24-DEC-1998 (first entry)
XX DE LKT-GnRH chimeric protein.
XX XX
XX KW Chimera; pCB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;
XX KW cytotoxic activity; antigen presentation; immune response; vaccine;
XX KW tumour.
XX OS Synthetic.
XX XX
XX PN WO9806848-A1.
XX XX

```


CC sequence which is highly immunogenic that can be used in the construction
 CC of a chimeric protein that comprises a leukotoxin polypeptide, several
 CC multimers, and the GnRH sequence. The chimeric protein can be used as a
 CC vaccine to help reduce the incidence of mammary tumours in a mammalian
 CC individual.

XX Sequence 49 AA;
 SQ Query Match 43.4%; Score 106; DB 19; Length 49;
 Best Local Similarity 50.0%; Pred. No. 1.5e-06;
 Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;
 QY 2 HWSYGLRPGSGPSLQYKANSKFITIGITELSSGSPSLHWSYGLRP 45
 |||||
 Db 2 HWSYGLRPGSGSQDWSY-----GLR--PGSSQHWSYGLRP 35
 |||||

RESULT 8
 AAW61542
 ID AAW61542 standard; Protein; 49 AA.
 XX AC AAW61542;
 XX DT 27-OCT-1998 (first entry)
 XX DE Peptide hormone GnRH-2 decapeptide (4 copies) fragment.
 XX KW GnRH; gonadotrophin releasing hormone peptide hormone; leukotoxin;
 KW immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;
 KW immune response; hormone receptor; cancerous cell; domestic animal;
 KW porcine; bovine; luteinizing hormone; follicle stimulating hormone;
 KW immunocastrate.
 XX OS Synthetic.
 XX PN WO9834639-A1.
 XX PD 13-AUG-1998.
 XX PF 04-FEB-1998; 98WO-CA00059.
 XX PR 05-FEB-1997; 97US-0036883.
 XX PA (BIOS-) BIOSTAR INC.
 XX PI Acres SD, Harland R, Manns JG;
 XX DR WPI: 1998-446952/38.
 XX DR N-PSDB; AAV45190.

PT Immunisation against endogenous molecules by administering vaccine
 PT to ear - useful to elicit efficient and uniform immune response
 PT against e.g. gonadotrophin releasing hormone to immunocastrate pigs
 and cattle

XX Example 1; Fig 1B; 6lpp; English.

XX This represents the amino acid sequence of the gonadotrophin releasing
 CC hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-
 CC GnRH polypeptide gene fusions. This is used to exemplify the method of
 CC invention of immunisation against endogenous molecules by administering
 CC a vaccine which comprises an immunogen and a carrier to the ear of the
 CC mammal. The method is useful for eliciting an efficient and uniform
 CC immune response to block or suppress the activity of an endogenous
 CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,
 CC or to elicit an immune response against a targeted endogenous cell type
 CC (e.g. a cancerous or otherwise diseased cell). It is especially useful
 CC to reduce the levels of GnRH in domestic animals, especially in porcine
 CC or bovine species. The use of GnRH immunogens in the vaccine reduces the
 CC levels of luteinizing hormone and follicle stimulating hormone and helps
 CC in immunocastrating the animal. Administration of vaccine compositions to
 CC the ear instead of intramuscular administration into the neck increases
 CC the efficiency of vaccination of mammals against endogenous immunogens,

CC and may increase uniformity of vaccine presentation since the ear is
 CC relatively uniform from animal to animal.

XX Sequence 49 AA;
 SQ Query Match 43.4%; Score 106; DB 19; Length 49;
 Best Local Similarity 50.0%; Pred. No. 1.5e-06;
 Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;
 QY 2 HWSYGLRPGSGPSLQYKANSKFITIGITELSSGSPSLHWSYGLRP 45
 |||||
 Db 2 HWSYGLRPGSGSQDWSY-----GLR--PGSSQHWSYGLRP 35
 |||||

RESULT 9
 AAY58363
 ID AAY58363 standard; Protein; 49 AA.
 XX AC AAY58363;
 XX DT 27-MAR-2000 (first entry)
 XX DE Four-copy gonadotrophin-releasing hormone (GnRH) multimer.
 XX KW GnRH multimer; gonadotrophin-releasing hormone; immunosterilisation;
 KW immunocastration; vaccine; feline; canine; equine; cervine; ds
 XX OS Mammalia.
 XX OS Synthetic.
 XX PN WO9962545-A2.
 XX PD 09-DEC-1999.
 XX PF 28-MAY-1999; 99WO-CA00493.
 XX PR 04-JUN-1998; 98US-0088024.
 XX PR 06-MAY-1999; 99US-0306689.
 XX PA (BIOS-) BIOSTAR INC.
 XX PI Robbins SC;
 XX DR WPI: 2000-086857/07.
 XX DR N-PSDB; AAZ55702.

PT Hormone immunogens, analogues or antibodies used to manufacture
 PT vaccines for suppression of reproductive behavior and fertility in
 PT vertebrates

XX Claim 5; Fig 5B; 80pp; English.

XX This sequence represents a four-copy gonadotrophin-releasing
 CC hormone (GnRH) multimer, where the second and fourth GnRH sequence
 CC have a His to Asp substitution at position 2 of the GnRH sequence.
 CC The invention relates to GnRH immunogens, analogues or antibodies
 CC that cross-react with endogenous GnRH of a vertebrate. A specifically
 CC claimed immunogenic fusion protein (AAY58361) comprises, in the N to
 CC C-terminal direction, a synthetic peptide sequence (AAY58364), an eight
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer
 CC protein), and a second eight copy GnRH multimer (which functions as a carrier
 CC be used in a vaccine composition for prepubertal administration to a
 CC vertebrate subject to result in prolonged suppression of reproductive
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are
 CC used to manufacture a composition or vaccine for immunosterilisation or
 CC immunocastration of feline, canine, equine or cervine subjects.
 CC The vaccines are used to suppress reproductive behaviour and/or
 CC fertility for at least 10 months. The prepubertal administration
 CC results in a prolonged, long-term suppression of testicular development
 CC and/or function in males, or a prolonged, long-term suppression of
 CC ovarian development and/or function in females. The methods provide a
 CC viable and desirable alternative to surgical forms of sterilisation that

CC invention relates to a method of using two GnRH immunogen vaccines to
 CC produce uncastrated male animals for meat production, one vaccination
 CC prior to or during the fattening period to reduce circulating
 CC testosterone levels, and the second vaccination about 2-8 weeks before
 CC slaughter to substantially reduce androgenic and/or non-androgenic
 CC steroids. The invention is used to produce food animals that exhibit the
 CC weight gain and muscle/fat distribution of male animals without the
 CC problems associated with male animals. Such problems include "boar
 CC taint", a urine-like odour found in cooked meat of uncastrated pigs which
 CC is caused by steroids stored in the tissues, and similar flavour
 CC impairments in the meat of other intact male animals. The invention is
 CC more reliable than prior art immunocastration techniques.

XX SQ Sequence 695 AA;

Query Match 44.5%; Score 108.5; DB 21; Length 695;
 Best Local Similarity 50.0%; Pred. No. 1.8e-05;
 Matches 22; Conservative 0; Mismatches 11; Indels 11; Gaps 1;

Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
 |||||
 Db 621 HWSYGLRPGSGQWSY-----GLRPGSQHWSYGLRP 653

RESULT 6
 AAW03944
 ID AAW03944 standard; Protein; 49 AA.

XX AC AAW03944;

XX DT 20-NOV-1996 (first entry)

XX DE GnRH 4-repeat sequence.

XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
 KW fusion protein; immunogen; vaccine; fertility control;
 KW contraceptive; sterilisation; PCB113; PCB11.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Peptide 1..10
 FT Peptide /label= GnRH
 FT Peptide 11..13
 FT Peptide /label= Spacer
 FT Peptide 14..23
 FT Peptide /label= GnRH
 FT Peptide 24..26
 FT Peptide /label= Spacer
 FT Peptide 27..36
 FT Peptide /label= GnRH
 FT Peptide 37..39
 FT Peptide /label= Spacer
 FT Peptide 40..49
 FT Peptide /label= GnRH

XX WO9624675-A1.

XX 15-AUG-1996.

XX PF 24-JAN-1996; 96WO-CA00049.

XX PR 10-FEB-1995; 95US-0387156.

XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Manns JG, Potter AA;

XX WPI; 1996-384447/38.

XX DR N-PSDB; AAT37178.

XX Gonadotropin-releasing hormone multimer fusion proteins - with
 PT leukotoxin in polypeptide for increased immunogenicity, useful in

PT antifertility vaccine prodn.
 XX
 XX Example 2; Fig 1B; 87pp; English.
 XX
 CC A synthetic DNA sequence (AAT37178) codes for a gonadotropin
 CC releasing hormone (GnRH) tetramer (AAW03944), in which the 4 GnRH
 CC repeat units are separated by spacers designed to increase
 CC immunogenicity. The DNA sequence was incorporated into vector
 CC pAA352 (ATCC 68283), which contains a truncated leukotoxin
 CC gene (LKT 352) derived from Pasteurella haemolytica, to
 CC give plasmid pCB113 (AAT37176). Escherichia coli transformants
 CC produce an LKT-GnRH fusion protein (see also AAW03942 and AAW03943)
 CC useful as a vaccine for fertility control of domestic or farm animals.

XX SQ Sequence 49 AA;

Query Match 43.4%; Score 106; DB 17; Length 49;
 Best Local Similarity 50.0%; Pred. No. 1.5e-06;
 Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
 |||||
 Db 2 HWSYGLRPGSGQWSY-----GLR--PGGSSQHWSYGLRP 35

RESULT 7

AAW79567

ID AAW79567 standard; Protein; 49 AA.

XX AC AAW79567;

XX DT 24-DEC-1998 (first entry)

XX DE GnRH-2.

XX Gonadotropin releasing hormone; GnRH; decapeptide; hypothalamus; spacer;
 KW LH; luteinising hormone; FSH; follicle stimulating hormone; vertebrate;
 KW pyroGlu; chimera; leukotoxin polypeptide; multimer; vaccine; tumour;
 KW immunogenic.

XX OS Synthetic.

XX WO9806848-A1.

XX PD 19-FEB-1998.

XX PF 08-AUG-1997; 97WO-CA00559.

XX PR 09-AUG-1996; 96US-0694865.

XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Manns JG, Potter AA;

XX WPI; 1998-159540/14.

XX DR N-PSDB; AAV61529.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -
 useful for, e.g. preparation of vaccines for reduction of incidence
 of mammary tumours in mammals

XX Disclosure; Figure 1B; 118pp; English.

XX The present sequence represents a recombinantly produced or chemically
 CC synthesised Gonadotropin releasing hormone-2 (GnRH-2) polypeptide, which
 CC contains four copies of the GnRH decapeptide (AAV61528) and triplet
 CC amino acid spacers between each of these sequences. This decapeptide is
 CC secreted naturally by the hypothalamus which controls release of both
 CC the luteinising hormone (LH) and the follicle stimulating hormone (FSH)
 CC in vertebrates. This sequence, as compared to the native peptide, has
 CC been found to have an N-terminal Gln rather than a pyroGlu residue, and
 CC also contains substitutions at amino acid residues 15 and 41, whereby His
 CC is replaced by Asp. This produces an alternating multimeric GnRH

Manns JG, Potter AA;
WPI; 1998-159540/14.
N-PSDB; AAV61535.

Chimeric protein of leukotoxin and gonadotropin releasing hormone -
useful for, e.g. preparation of vaccines for reduction of incidence
of mammary tumours in mammals

Claim 9; Figure 9.1-6; 118pp; English.

The present sequence represents the LKT-GnRH chimeric protein from
pcB122. This plasmid contains the LKT lll polypeptide fused to sixteen
copies of the GnRH peptide, in the pattern of: 8 copies of GnRH-LKT lll-8
copies of GnRH. This chimera lacks cytotoxic activity which enables
there to be an increase in antigen presentation and thus an optimal
immune response. The removal of this region also enables the truncated
LKT to be expressed at much higher levels and allows the amount of
antigen administered to be reduced. This chimeric protein comprises a
leukotoxin polypeptide, several multimers, and a GnRH sequence. The
chimeric protein can be used as a vaccine to help reduce the incidence of
mammary tumours in a mammalian individual.

Sequence 695 AA;

Query Match 44.5%; Score 108.5; DB 19; Length 695;
Best Local Similarity 50.0%; Pred. No. 1.8e-05;
Matches 22; Conservative 0; Mismatches 11; Indels 11; Gaps 1;

QY 2 HWSYGRLPGSGSLQYIKANSKFIGITELSSGPSLHWSYGLRP 45
||||||| | | | | | | |||||
Db 621 HWSYGRLPGSQDWSY-----GLRFGGSQHWSYGLRP 653

RESULT 4
AAV58361
ID AAY58361 standard; Protein; 695 AA.
XX AC AAY58361;
XX DT 27-MAR-2000 (first entry)
XX DE Leukotoxin/gonadotropin-releasing hormone fusion protein.
XX KW Leukotoxin; gonadotropin-releasing hormone; GnRH; immunosterilisation;
KW KWK immunocontraception; vaccine; feline; canine; equine; cervine.
OS Chimeric - Pasteurella haemolytica.
OS OS Chimeric - Mammalia.
XX PN WO9962545-A2.
XX PD 09-DEC-1999.
XX PF 28-MAY-1999; 99WO-CA00493.
XX PR 04-JUN-1998; 98US-0088024.
XX PR 06-MAY-1999; 99US-0306689.
XX XX (BIOS-) BIOSTAR INC.
PA Robbins SC;
XX PI Robbins SC;
XX PI WPI: 2000-086857/07.
DR N-PSDB; AA255700.
XX Hormone immunogens, analogues or antibodies used to manufacture
PT vaccines for suppression of reproductive behavior and fertility in
PT vertebrates -
XX Claim 20; Fig 6A-6F; 88pp; English.
XX PS This sequence represents a fusion protein comprising gonadotropin-
CC

releasing hormone (GnRH) immunogens and a Pasteurella
haemolytica leukotoxin (LKT) protein. The fusion protein comprises, in
the N to C-terminal direction, a synthetic peptide sequence (AAV58364),
an eight copy GnRH multimer (composed of two copies of the 4xGnRH
multimer sequence of AAV58363), the LKT protein (which functions as a
carrier protein), and a second eight copy GnRH multimer. The fusion
protein may be used in a vaccine composition for prepubertal
administration to a vertebrate subject to result in prolonged suppression
of reproductive behaviour and/or fertility. GnRH immunogens, analogues or
antibodies that cross-react with endogenous GnRH of a vertebrate subject
are used to manufacture a composition or vaccine for immunosterilisation
or immunocontraception of feline, canine, equine or cervine subjects.
The vaccines are used to suppress reproductive behaviour and/or
fertility for at least 10 months. The prepubertal administration
results in a prolonged, long-term suppression of testicular development
and/or function in males, or a prolonged, long-term depression of
ovarian development and/or function in females. The methods provide a
viable and desirable alternative to surgical forms of sterilisation that
are currently used.

Sequence 695 AA;

Query Match 44.5%; Score 108.5; DB 21; Length 695;
Best Local Similarity 50.0%; Pred. No. 1.8e-05;
Matches 22; Conservative 0; Mismatches 11; Indels 11; Gaps 1;

QY 2 HWSYGRLPGSGSLQYIKANSKFIGITELSSGPSLHWSYGLRP 45
||||||| | | | | | | |||||
Db 621 HWSYGRLPGSQDWSY-----GLRFGGSQHWSYGLRP 653

RESULT 5
AAV58133
ID AAY58133 standard; Protein; 695 AA.
XX AC AAY58133;
XX DT 07-MAR-2000 (first entry)
XX DE Gonadotropin releasing hormone-leukotoxin fusion protein.
XX KW Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
non-androgenic; steroid; reduction; weight gain; muscle distribution;
fat distribution; male pattern; boar taint; flavour; impairment;
reliable; immunocastration; meat production.
XX OS Chimeric - Mammalia.
OS OS Chimeric - Pasteurella haemolytica.
XX PN WO9956771-A2.
XX PD 11-NOV-1999.
XX PF 05-MAY-1999; 99WO-CA00360.
XX PR 05-MAY-1998; 98US-0084217.
XX PA (BIOS-) BIOSTAR INC.
XX PI Manns JG, Acres SD, Harland R;
XX XN WPI: 2000-062125/05.
DR N-PSDB; AA246400.
XX Production of uncastrated male food animals using vaccines -
PT Claim 22; Fig 3A-3F; 87pp; English.
XX PS This sequence represents a chimeric gonadotropin releasing
CC hormone (GnRH)-leukotoxin (LKT) fusion protein, which may be
used as a vaccine. The LKT portion of the protein acts to enhance
CC the immunogenicity of the multimeric GnRH portion (AAV58135). The


```

PA (LADD/) LADD A E.
PA (WANG/) WANG C Y.
XX (ZAMB/) ZAMB T.
PI Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
DR Immunogenic luteinising hormone releasing hormone peptide(s) -
XX that suppress LHRH activity in males and females
PT Claim 8; Page 88; 213pp; English.
PS
XX Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The
CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
CC which elicits an immune response to the coupled peptide in members of
CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
CC an adjuvant peptide sequence from the invasive protein of Yersinia.
CC Invasin and Th domains and between the immune stimulator and hapten
CC components. When the hapten is LHRH, then optionally the invasive domain
CC can be omitted from the immune stimulator component.
CC The present sequence represents an LHRH-containing immunogenic peptide
CC as above which can be used as a potent vaccine for treating e.g.
CC prostatic hyperplasia, androgen-dependent carcinoma, prostatic
CC carcinoma, testicular carcinoma, endometriosis, benign uterine tumours,
CC recurrent functional ovarian cysts, (severe) premenstrual syndrome or
CC oestrogen-dependent breast cancer, or for induction of infertility.
XX
SQ Sequence 47 AA;
Query Match 46.9%; Score 114.5; DB 15; Length 47;
Best Local Similarity 72.7%; Pred. No. 1.1e-07;
Matches 24; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 13 GPSLQYIKANSKFIGITELSSGSLHWSYGLRP 45
Db 17 GKKQYIKANSKFIGITELGGE--HWSYGLRP 46

RESULT 2
AAR62701
ID AAR62701 standard; peptide; 27 AA.
XX
AC AAR62701;
XX
DT 10-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
DE Helper T cell epitope; universal immune stimulator; invasive; hapten;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility;
KW tetanus toxin.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1..17
FT /note= "tetanus toxin helper T cell epitope"
FT Domain 18..27
FT /note= "LHRH hapten"
XX
PN W09425060-A.
XX
PD 10-NOV-1994.
XX
PF 28 APR-1994; 94WO-US04832.
XX
PR 27-APR-1993; 93US-0057166.

```

```

PR 14-APR-1994; 94US-0229275.
XX (LADD/) LADD A E.
PA (WANG/) WANG C Y.
PA (ZAMB/) ZAMB T.
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
DR Immunogenic luteinising hormone releasing hormone peptide(s) -
XX that suppress LHRH activity in males and females
PT Claims 8, 12; Page 84; 213pp; English.
PS
XX Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The
CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
CC which elicits an immune response to the coupled peptide in members of
CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
CC an adjuvant peptide sequence from the invasive protein of Yersinia.
CC Invasin and Th domains and between the immune stimulator and hapten
CC components. When the hapten is LHRH, then optionally the invasive domain
CC can be omitted from the immune stimulator component.
CC The present sequence represents an LHRH-containing, invasive-free
CC immunogenic peptide as above which can be used as a potent vaccine for
CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
CC uterine tumours, recurrent functional ovarian cysts, (severe)
CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
CC induction of infertility.
XX
SQ Sequence 27 AA;
Query Match 46.5%; Score 113.5; DB 15; Length 27;
Best Local Similarity 79.3%; Pred. No. 7.6e-08;
Matches 23; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 17 QYIKANSKFIGITELSSGSLHWSYGLRP 45
Db 3 QYIKANSKFIGITEL-----HWSYGLRP 26

RESULT 3
AAW79573
ID AAW79573 standard; Protein; 695 AA.
XX
AC AAW79573;
XX
DT 24-DEC-1998 (first entry)
XX
DE LKT-GnRH chimeric protein.
XX
KW Chimera; pCB122; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
KW cytotoxic activity; antigen presentation; immune response; vaccine;
KW tumour.
XX
OS Synthetic.
XX
PN W09806848-A1.
XX
PD 19-FEB-1998.
XX
PF 08-AUG-1997; 97WO-CA00559.
XX
PR 09-AUG-1996; 96US-0694865.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 24.1795 Seconds
(without alignments)
211.311 Million cell updates/sec

Title: US-09-848-834A-19

Perfect score: 244

Sequence: 1 XHWSYGLRPGSGPSLQYIK.....GTELSSGSPSLHWSYGLRFX 46

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114.5	46.9	47	15	AA62723
2	113.5	46.5	27	15	AA62701
3	108.5	44.5	695	19	AAW79573
4	108.5	44.5	695	21	AAV58361
5	108.5	44.5	695	21	AAV58133
6	106	43.4	49	17	AAW03944
7	106	43.4	49	19	AAW79567
8	106	43.4	49	19	AAW61542
9	106	43.4	49	21	AAV58363
10	106	43.4	49	21	AAV58135
11	106	43.4	544	17	AAW03943

12	106	43.4	544	19	AAW79570	LKT-GnRH chimeric
13	106	43.4	977	17	AAW03942	LKT-GnRH protein f
14	106	43.4	977	19	AAW79569	LKT-GnRH chimeric
15	94	38.5	42	21	AA620865	GnRH tandem dimer
16	92.5	37.9	23	21	AA620864	GnRH tandem repeat
17	91	37.3	30	11	AA607323	Luteinising hormon
18	91	37.3	40	20	AA631183	Ubiquitin fusion p
19	91	37.3	40	22	AA671954	GnRH sequence #2.
20	91	37.3	41	20	AA631182	Ubiquitin fusion p
21	91	37.3	41	22	AA671953	GnRH sequence #1.
22	90	36.9	40	21	AA696085	Cattle gonadotropi
23	90	36.9	283	12	AA611186	Plasmid pBTA862-en
24	90	36.9	323	12	AA611187	Plasmid pBTA859-en
25	90	36.9	398	21	AA696090	BHV-1 truncated gd
26	90	36.9	399	21	AA696093	BHV-1 truncated gd
27	90	36.9	411	21	AA696089	GnRH tetramer-BHV-
28	90	36.9	442	21	AA696091	GnRH tetramer-BHV-
29	89	36.5	216	21	AA692665	MUC-1 analogue con
30	89	36.5	750	21	AA692638	Mutant human prost
31	87	35.7	20	19	AA647438	Antigenic peptide.
32	87	35.7	20	20	AA631174	Ubiquitin fusion p
33	87	35.7	20	20	AA631177	Ubiquitin fusion p
34	87	35.7	20	20	AA631178	Ubiquitin fusion p
35	87	35.7	20	20	AA631179	Ubiquitin fusion p
36	87	35.7	20	22	AA671945	GnRH dimer. unde
37	87	35.7	20	22	AA671948	GnRH dimer peptide
38	87	35.7	20	22	AA671949	GnRH mixed dimer p
39	87	35.7	20	22	AA671950	GnRH mixed dimer p
40	87	35.7	21	11	AA607324	Luteinising hormon
41	87	35.7	263	12	AA611185	Plasmid pBTA870-en
42	85	34.8	109	22	AA620147	Growth differentia
43	84	34.4	116	21	AA645502	Modified murine in
44	84	34.4	116	21	AA645526	Modified murine in
45	83.5	34.2	158	19	AAW81327	TNF-1, a TNF-alpha

ALIGNMENTS

RESULT 1
AA62723
ID AA62723 standard; peptide; 47 AA.
XX
AC AA62723;
XX
DT 17-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
KW Helper T cell epitope; universal immune stimulator; invasin; haptan;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility;
KW tetanus toxin.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Domain 1..16 /note= "invasin domain"
FT Domain 19..35 /note= "tetanus toxin helper T cell epitope"
FT Domain 38..47 /note= "LHRH haptan"

FT WO9425060-A.
XX
PD 10-NOV-1994.
XX
PF 28-APR-1994; 94WO-US04832.
XX
PR 27-APR-1993; 93US-0057166.
PR 14-APR-1994; 94US-0229275.
XX


```

RX MEDLINE=211396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Gallibert F., Guzy J.,
RA Gurjal M., Hong A., Huizar L., Hymen R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium melliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
RW EMBL: AE007266; AAK65455.1;
KW Isomerase; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 374 AA; 40999 MW; 0369AA67790B23D6 CRC64;

Query Match 24.0%; Score 58.5; DB 16; Length 374;
Best Local Similarity 41.2%; Pred. No. 13;
Matches 14; Conservative 6; Mismatches 13; Indels 1; Gaps 1;

QY 2 HWSYGLRPGSSGSPSYQYKANSKFGITELSSGP 35
Db |||||:| | | | | | | | | | | | | | | |
21 HWSYGLRPGSSGSPSYQYKANSKFGITELSSGP 35

RESULT 14
QXID7 PRELIMINARY; PRT; 530 AA.
AC Q9XID7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE HYPOTHETICAL 59.7 KDA PROTEIN F23M19.1 (LATE EMBRYOGENESIS ABUNDANT
DE PROTEIN, PUTATIVE).
GN F23M19.1 OR F7P12.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J., Liu A., Li J., Kremenetskaia I., Luros J., Gonzalez A.,
RA Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,
RA Hansen N., Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F23M19 sequence.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Theologis A.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altai H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egtu P., Feldblyum T.V., Feng J.D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis

```

```

RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL: AC007454; AAD39602.1;
DR EMBL: AC023913; AAG51902.1;
DR InterPro: IPR000379; Est_lip_thioest_actsite.
KW Hypothetical protein.
SQ SEQUENCE 530 AA; 59670 MW; EA07EC8C787A589 CRC64;

Query Match 23.4%; Score 57; DB 10; Length 530;
Best Local Similarity 37.9%; Pred. No. 31;
Matches 11; Conservative 9; Mismatches 7; Indels 2; Gaps 1;

QY 10 GSSGSPLOYKANKSKFGITELSSGSPSLH 38
Db | | | | | | | | | | | | | | | |
395 GLTGSSLLWVVRATNEFLGV--LSCSPYMH 421

RESULT 15
Q989H0 PRELIMINARY; PRT; 289 AA.
ID Q989H0;
AC Q989H0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE MLL6426 PROTEIN.
DE MLL6426.
GN MLL6426.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
RR EMBL: AP003009; BAB52726.1;
DR InterPro: IPR003265; Endo_3c.
DR InterPro: IPR003383; HHH_1.
DR SMART: SM00478; ENDO3c; 1.
DR SMART: SM00278; Hhh1; 1.
KW Complete proteome.
SQ SEQUENCE 289 AA; 30929 MW; 14A288A1784AC98A CRC64;

Query Match 22.5%; Score 55; DB 16; Length 289;
Best Local Similarity 29.8%; Pred. No. 28;
Matches 17; Conservative 5; Mismatches 15; Indels 20; Gaps 2;

QY 6 GLRPGSSGSPSYQYKANSKFGITELSSGSPSLH 45
Db | | | | | | | | | | | | | | | |
209 GTFGGATPASDYVSLESFICFCRAAGLRASLDALMWRTMRISIGPSF---AGLRP 262

```

Search completed: October 10, 2002, 16:10:09
Job time : 19.8889 secs

RT "Expression of two gonadotropin-releasing hormone (GnRH) precursor
 RL genes in various tissues of the Japanese eel and evolution of GnRH."
 Zool. Sci. 16:471-478(1999).
 [2]

RA Okubo K., Suetake H., Aida K.:
 RN SEQUENCE FROM N.A.
 RT "A splicing variant for the prepro-mammalian gonadotropin-releasing
 RT hormone (prepro-mGnRH) mRNA is present in the brain and various
 RT peripheral tissues of the Japanese eel."
 Zool. Sci. 16:645-651(1999).
 RL Zool. Sci. 16:645-651(1999).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.

DR EMBL; AB026989; BAA82608.1; -.

DR EMBL; AB026991; BAA83597.1; -.

DR InterPro; IPR002012; GnRH.

DR InterPro; IPR004079; Gonadoliberin.

DR Pfam; PF00446; GnRH; 1.

DR PRINTS; PR01541; GONADOLIBRN1.

DR PROSITE; PS00473; GnRH; 1.

KW Annotation: Hormone; Signal.

FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 32 MGNRH.

FT CHAIN 33 91 GnRH ASSOCIATED PEPTIDE.

SQ SEQUENCE 91 AA; 9893 MW; BAI5C9DC08434A7B CRC64.

Query Match 24.2%; Score 59; DB 13; Length 91;

Best Local Similarity 66.7%; Pred. No. 2.2;

Matches 12; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLQ 17

|||||

DB 24 HWSYGLRPGGKRGADSLQ 41

RESULT 10

ID Q92201 PRELIMINARY; PRT; 385 AA.

AC Q92201;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DE 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)

DE FIBRINOGEN A-ALPHA CHAIN (FRAGMENT).

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.

OX NCBI_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

RA Murakawa M.;

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF095466; AAC67565.1; -.

DR HSSP; P02671; 1PZF.

FT NON_TER 1 385

FT NON_TER 385 385

SQ SEQUENCE 385 AA; 40909 MW; FA576CA0E2A2101A CRC64;

Query Match 24.2%; Score 59; DB 11; Length 385;

Best Local Similarity 38.1%; Pred. No. 12;

Matches 16; Conservative 2; Mismatches 12; Indels 12; Gaps 2;

QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSPSLHWSYGL 43

|||||

DB 128 HWI-----PGSSGPGSDQSPVSS-----PSGSGSHWSSGV 157

RESULT 11

Q96SA1

ID Q96SA1 PRELIMINARY; PRT; 329 AA.

AC Q96SA1;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DE PUTATIVE MUCONATE CYCLOISOMERASE (EC 5.5.1.1).

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

DE TNFAIP1-LIKE PROTEIN.

GN FKSG86.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Wang Y.-G., Gong L.;

RT "Cloning and characterization of FKSG86, a novel gene encoding a

RT TNFAIP1-like protein."

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY027918; AAK27301.1; -.

SQ SEQUENCE 329 AA; 36371 MW; 61092C2B0E81EA2D CRC64;

Query Match 24.0%; Score 58.5; DB 4; Length 329;

Best Local Similarity 37.1%; Pred. No. 11;

Matches 13; Conservative 9; Mismatches 10; Indels 3; Gaps 1;

QY 6 GLRPGSGPSLQYIKANSKFIGITELSSGSPSLHWS 40

|||||

DB 23 GLRPGPAAYGLKPLTPNSKYV---KLVNKGSLHYT 54

RESULT 12

Q96P93

ID Q96P93 PRELIMINARY; PRT; 329 AA.

AC Q96P93;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DE POLYMERASE DELTA-INTERACTING PROTEIN 1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=21477421; PubMed=11593007;

RT He H., Tan C.K., Downey K.M., So A.G.;

RT "A tumor necrosis factor alpha - and interleukin 6-inducible protein

RT that interacts with the small subunit of DNA polymerase delta and

RT proliferating cell nuclear antigen."

RL Proc. Natl. Acad. Sci. U.S.A. 98:11979-11984(2001).

DR EMBL; AF401315; AAL14962.1; -.

SQ SEQUENCE 329 AA; 36395 MW; D356490B48995187 CRC64;

Query Match 24.0%; Score 58.5; DB 4; Length 329;

Best Local Similarity 37.1%; Pred. No. 11;

Matches 13; Conservative 9; Mismatches 10; Indels 3; Gaps 1;

QY 6 GLRPGSGPSLQYIKANSKFIGITELSSGSPSLHWS 40

|||||

DB 23 GLRPGPAAYGLKPLTPNSKYV---KLVNKGSLHYT 54

RESULT 13

Q92YR6

ID Q92YR6 PRELIMINARY; PRT; 374 AA.

AC Q92YR6;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DE PUTATIVE MUCONATE CYCLOISOMERASE (EC 5.5.1.1).

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

[illegible]

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 17.8889 seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-19
Perfect score: 244
Sequence: 1 XHWSYGLRPGSSGSLQVIK.....GITELSSGSLHWSYGLRXPX 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_oranelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	31.1	1310	2 Q93N27	Q93n27 clostridium
2	71	29.1	1331	10 Q9SF03	Q9sf03 arabidopsis
3	64	26.2	82	13 Q90VY3	Q90vy3 oncorhynch
4	63	25.8	82	13 Q918P9	Q918p9 oncorhynch
5	61	25.0	458	16 Q9A409	Q9a409 caulobacter
6	60	24.6	447	16 Q9A409	Q9a409 pseudomonas
7	59.5	24.4	90	13 Q9YF63	Q90yf63 rana catesb
8	59.5	24.4	1172	4 Q9Y4F2	Q9y4f2 homo sapien
9	59	24.2	91	13 Q9PRH0	Q9prh0 anguilla ja
10	59	24.2	385	11 Q9Z201	Q9z201 cavia porce
11	58.5	24.0	329	4 Q96SA1	Q96sa1 homo sapien
12	58.5	24.0	329	4 Q96P93	Q96p93 homo sapien
13	58.5	24.0	374	16 Q92YR6	Q92yr6 rhizobium m
14	57	23.4	530	10 Q9XID7	Q9xid7 arabidopsis
15	55	22.5	289	16 Q989H0	Q989h0 rhizobium l
16	55	22.5	364	10 Q9LUE6	Q9lue6 arabidopsis

17	55	22.5	417	16 Q9PGX9	Q9pgx9 xylella fas
18	55	22.5	871	5 Q44358	Q44358 nephila cla
19	55	22.5	2249	5 Q9NHW4	Q9nhw4 nephila cla
20	54.5	22.3	435	12 Q9WPM1	Q9wpm1 blueetongue
21	54	22.1	104	11 Q9JL82	Q9j182 mus musculus
22	54	22.1	324	8 Q9G6T7	Q9g6t7 sardinops m
23	54	22.1	324	8 Q94WE2	Q94we2 gobiosoma m
24	54	22.1	324	8 Q94WD6	Q94wd6 bathygobius
25	54	22.1	1074	10 Q9SYZ6	Q9syz6 arabidopsis
26	53.5	21.9	3944	5 Q18667	Q18667 caenorhabdi
27	53	21.7	82	13 Q918Q0	Q918q0 oncorhynch
28	53	21.7	265	2 Q9EWX0	Q9ewx0 streptomyce
29	53	21.7	306	10 Q9SDN3	Q9sdn3 prunus dulc
30	53	21.7	417	2 Q9F7Q8	Q9f7q8 uncultured
31	53	21.7	444	16 P71718	P71718 mycobacteri
32	53	21.7	539	11 Q925P4	Q925f4 mus musculu
33	53	21.7	1146	13 Q90584	Q90584 gallus gall
34	53	21.7	1494	11 Q88902	Q88902 rattus norv
35	52.5	21.5	146	17 Q27851	Q27851 methanother
36	52.5	21.5	221	2 Q9KW73	Q9kw73 staphylococ
37	52.5	21.5	230	4 Q9BX33	Q9bx33 homo sapien
38	52.5	21.5	464	11 Q61965	Q61965 mus musculu
39	52.5	21.5	475	16 Q9AB89	Q9ab89 caulobacter
40	52.5	21.5	3112	5 Q9NKP1	Q9nkp1 leishmania
41	52	21.3	62	13 Q90ZE1	Q90ze1 oncorhynch
42	52	21.3	82	13 Q92094	Q92094 oncorhynch
43	52	21.3	82	13 Q9W7G1	Q9w7g1 oncorhynch
44	52	21.3	87	13 Q9YI26	Q9yi26 sparus aura
45	52	21.3	88	13 Q9PSY9	Q9psy9 sparus aura

ALIGNMENTS

RESULT 1
Q93N27
ID Q93N27 PRELIMINARY; PRT; 1310 AA.
AC Q93N27;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE TETANUS TOXIN (FRAGMENT).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA Shumin Z., Dianliang L.;
RT "Cloning and sequence analysis of tetanus toxin gene."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389424; AAK72964.2; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;
Query Match 31.1%; Score 76; DB 2; Length 1310;
Best Local Similarity 93.8%; Pred. No. 0.26;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 16 LOYKANSKFIGITEL 31
:|||||
Db 830 MOYKANSKFIGITEL 845
:
RESULT 2
Q9SF03
ID Q9SF03 PRELIMINARY; PRT; 1331 AA.
AC Q9SF03;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE F26K24.25 PROTEIN (GENOMIC DNA, CHROMOSOME 3, P1 CLONE: MEC18).

CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U63326; AAB16837.1; -.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBRN1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 92 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 33 33 SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
Query Match 23.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
|||||||

Search completed: October 10, 2002, 16:06:38
Job time : 6.99573 secs

```
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pharmaceutical; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 GNRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT CONFLICT 16 16 W -> S (IN REF. 3).
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 23.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
Db |||||
25 HWSYGLRPG 33

RESULT 14
GONL_RAT STANDARD; PRT; 92 AA.
ID GONL_RAT
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); Prolactin release-inhibiting factor
DE I].
DE GNRH1 OR GNRH.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RX MEDLINE=Heart;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus.";
RL Science 235:1514-1517(1987).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.

-!- SUBCELLULAR LOCATION: Secreted.
-!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
-!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; S50870; AAB24572.1; -.
DR EMBL; M12579; AAA41263.1; -.
DR EMBL; M31670; AAA41264.1; -.
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -.
DR EMBL; M15528; -. NOT_ANNOTATED_CDS.
DR PIR; B26173; RHRTG
DR PIR; A48410; A48410.
DR InterPro; IPR002012; Gnrh.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; Gnrh; 1.
DR PRINTS; PR01541; GONADOLIBRNI.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10500 MW; 49AB5C64DA8A3EB3 CRC64;

Query Match 23.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPG 10
Db |||||
25 HWSYGLRPG 33

RESULT 15
GONL_TUPGB STANDARD; PRT; 92 AA.
ID GONL_TUPGB
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Prognadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I].
DE GNRH1 OR GNRH.
GN Tupia glis belangeri (Common tree shrew).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=9396;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGnrh mRNAs in the tree shrew:
RT first direct evidence for mesencephalic Gnrh gene expression in a
RT placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
```

GN GNRH1 OR GNRH.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RA Weesner G.D., Matteri R.L., Becker B.A.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=72114303; PubMed=4946067;
 RA Baba Y., Matsuo H., Schally A.V.;
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.
 RT Confirmation of the proposed structure by conventional sequential
 RT analyses";
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
 RN [3]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72065376; PubMed=4942726;
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 RT phase method";
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
 RN [4]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE=72117544; PubMed=4946275;
 RA Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L32864; AAA31066.1; -;
 DR PIR; A01411; RHFGG.
 DR InterPro; IPR002012; GNRH.
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBRN.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1
 FT CHAIN 1 23
 FT PEPTIDE 24 91 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
 Query Match 23.8%; Score 58; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.77;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPG 10
 Db 25 HWSYGLRPG 33

RESULT 13
 GONL_HUMAN
 ID GONL_HUMAN STANDARD; PRT; 92 AA.
 AC P01148;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GNRH I) (Luliberin I) (Gonadorelin); GNRH-associated
 DE peptide I].
 OS Homo sapiens (Human).
 GN GNRH1 OR GNRH OR LHRH.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89366682; PubMed=2671939;
 RA Haylick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 RT hormone gene";
 RL Nucleic Acids Res. 17:6403-6403(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Haylick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85012739; PubMed=6050951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 RT releasing hormone.";
 RL Nature 311:666-668(1984).
 RN [4]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=83126573; PubMed=6760865;
 RA Tan L., Rousseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide
 RT biosynthesized in the human placenta";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 CC Lutrepulse or Lutrel (Ferring Pharmaceuticals) and Relisorm
 CC (Serono).
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X01059; CAA25526.1; -;
 DR EMBL; M12578; AAA35916.1; -;
 DR EMBL; X15215; CAA33285.1; -;
 DR PIR; A01410; RHUG.
 DR PIR; A26173; A26173.
 DR PIR; S05308; S05308.
 DR MIM; 152760;
 DR InterPro; IPR002012; GNRH.
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBRN.

RT spectrometry-decapeptide-Edman degradation).";
 RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U02517; AAA03433.1; -
 DR PIR; A93780; RHSHG.
 DR InterPro; IPR002012; GNRH.
 DR Pfam; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta.
 KW NON_TER 1 1
 FT CHAIN 1 >61 PROGONADOLIBERIN I.
 FT PEPTIDE 1 10 GONADOLIBERIN I.
 FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
 FT NON_TER 61 61
 SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;
 Query Match 24.6%; Score 60; DB 1; Length 61;
 Best Local Similarity 52.2%; Pred. No. 0.27;
 Matches 12; Conservative 1; Mismatches 2; Indels 8; Gaps 1;
 YQ 2 HWSYGLRPGSSGPSLOYIKANSK 24
 Db |||||
 |||||
 2 HWSYGLRPGG-----KRNK 16
 RESULT 8
 ID VG84_BPML5 STANDARD; PRT; 66 AA.
 AC Q05301;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DE 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Gene 84 protein (GP84).
 GN 84.
 OS Mycobacteriophage L5.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxID=31757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93211282; PubMed=8459766;
 RA Hatfull G.F., Sarkis G.J.;
 RT "DNA sequence, structure and gene expression of mycobacteriophage L5:
 RT a phage system for mycobacterial genetics.";
 RL Mol. Microbiol. 7:395-405(1993).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z18946; CAA79460.1; -
 DR PIR; S31029; S31029.

SQ SEQUENCE 66 AA; 7424 MW; 9C7104C7A4FA74A5 CRC64;
 Query Match 24.4%; Score 59.5; DB 1; Length 66;
 Best Local Similarity 45.5%; Pred. No. 0.34;
 Matches 15; Conservative 2; Mismatches 7; Indels 9; Gaps 2;
 YQ 5 YGL-----RPGSSGPSLOYIKANSKFIGITELSS 32
 Db |||||
 |||||
 36 YGFEVDWYEPGESG-----YIKRNGKFGVTWEVS 64
 RESULT 9
 ID GONL_MOUSE STANDARD; PRT; 90 AA.
 AC P13562;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Conadotropin releasing
 DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor
 DE I].
 GN GNRH1 OR GNRH.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87069928; PubMed=3024317;
 RA Mason A.J., Haylick J.S., Zoeller R.T., Young W.S. III,
 RA Phillips H.S., Nikolics K., Seeburg P.H.;
 RT "A deletion truncating the gonadotropin-releasing hormone gene is
 RT responsible for hypogonadism in the hpg mouse.";
 RL Science 234:1366-1371(1986)
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M14872; AAA37717.1; -
 DR MGD; MGI:95789; Gnrh.
 DR InterPro; IPR002012; GNRH.
 DR InterPro; IPR004079; GonadolibereinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBERNI.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 KW SIGNAL 1 21
 FT CHAIN 22 90 PROGONADOLIBERIN I.
 FT PEPTIDE 22 31 GONADOLIBERIN I.
 FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
 FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826B4D9 CRC64;
 Query Match 24.0%; Score 58.5; DB 1; Length 90;
 Best Local Similarity 27.9%; Pred. No. 0.65;
 Matches 12; Conservative 9; Mismatches 17; Indels 5; Gaps 1;

YQ 2 HWSYGLRPGSSGPSLOYIKANSKFIGITELSSGP-----SLHW 39

```

RN  SEQUENCE FROM N.A.
RP  TISSUE=Muscle;
RA  Klungland H., Anderson O., Alestroem P.;
RT  "The salmon gonadotropin-releasing hormone encoding gene in
RT  salmonids.";
RL  Mol. Mar. Biol. Biotechnol. 1:420-425(1992).
[2]
RN  SEQUENCE OF 1-65 FROM N.A.
RX  MEDLINE=93386322; PubMed=1308825;
RA  Alestroem P., Klsen G., Klungland H., Andersen O.;
RT  "Fish gonadotropin-releasing hormone gene and molecular approaches
RT  for control of sexual maturation: development of a transgenic fish
RT  model.";
RL  Mol. Mar. Biol. Biotechnol. 1:376-379(1992).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
DR  EMBL; X79710; CAA56149.1; -.
DR  EMBL; S65569; AAD13966.1; -.
DR  InterPro: IPR002012; GnRH.
DR  Pfam: PF00446; GnRH; 1.
DR  PROSITE; PS00473; GnRH; 1.
KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW  Signal.
FT  NON_TER 1 1
FT SIGNAL <1 15 BY SIMILARITY.
FT CHAIN 16 74 PROGNADOLIBERIN III.
FT PEPTIDE 16 25 GONADOLIBERIN III.
FT PEPTIDE 29 74 GnRH-ASSOCIATED PEPTIDE III.
FT MOD_RES 16 16 PYRROLIDONE CARBOXYLIC ACID
FT FT (BY SIMILARITY).
FT MOD_RES 25 25 AMIDATION (G-26 PROVIDE AMIDE GROUP).
FT SEQUENCE 74 AA; 8254 MW; BD63C46D8228EF84 CRC64;
SQ
Query Match 26.2%; Score 64; DB 1; Length 74;
Best Local Similarity 38.3%; Pred. No. 0.1;
Matches 18; Conservative 5; Mismatches 20; Indels 4; Gaps 2;
QY 2 HWSYGLRPGSGPSLOYIKANSKFI---GITELSSGSPSLHWSYGLRP 45
Db 17 HWSYGLPGGK-RSVGELEATIKMWDGTVVLPETSAHVSRILRP 62
RESULT 6
GONI_MESAU
ID GONI_MESAU STANDARD; PRT; 63 AA.
AC 009163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
DE GN GNRI OR GNRI OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
```

```

RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR  EMBL; U91938; AAB51302.1; -.
DR  InterPro: IPR002012; GnRH.
DR  Pfam: PF00446; GnRH; 1.
DR  PROSITE; PS00473; GnRH; 1.
KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW  Placenta.
FT  NON_TER 1 1
FT CHAIN 1 >63 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >63 GnRH-ASSOCIATED PEPTIDE I (BY
FT FT SIMILARITY).
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT FT ACTIVITY (BY SIMILARITY).
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
FT FT SIMILARITY).
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
FT FT SIMILARITY).
FT NON_TER 63 63
FT SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;
SQ
Query Match 25.6%; Score 62.5; DB 1; Length 63;
Best Local Similarity 31.4%; Pred. No. 0.13;
Matches 16; Conservative 6; Mismatches 8; Indels 21; Gaps 3;
QY 2 HWSYGLRPGSGPSLOYIKANSKFI-----ITELSSGSPSL----HW 39
Db 2 HWSYGLRPGG-----KRNAERLGDSFQEMKDEVDQAEPOHLECTVHW 44
RESULT 7
GONI_SHEEP
ID GONI_SHEEP STANDARD; PRT; 61 AA.
AC Q28588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
DE (luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
DE GN GNRI OR GNRI OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RA STRAIN=WESTERN RANGE; TISSUE=Hypothalamus;
RA Rodriguez R.E., Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-10.
RX MEDLINE=72094314; PubMed=4550508;
RA Burgos R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vale W., Guillemin R.;
RA "Primary structure of the ovine hypothalamic luteinizing hormone-
RA releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 5.99573 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-19

Perfect score: 244

Sequence: 1 XHWSYGLRPGSGPSLQYIK.....GITELSSGSLHWSYGLRPX 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	31.1	1314	1	TETX_CLOTE
2	69	28.3	82	1	GON3_SALTR
3	65	26.6	74	1	GON3_ONCTS
4	65	26.6	82	1	GON3_SALSA
5	64	26.2	74	1	GON3_ONCMY
6	62.5	25.6	63	1	GON1_MESAU
7	60	24.6	61	1	GON1_SHEEP
8	59.5	24.4	66	1	VG84_BPML5
9	58.5	24.0	90	1	GON1_MOUSE
10	58	23.8	67	1	GON1_MACMU
11	58	23.8	89	1	GON1_XENLA
12	58	23.8	91	1	GON1_PIG
13	58	23.8	92	1	GON1_HUMAN
14	58	23.8	92	1	GON1_RAT
15	58	23.8	92	1	GON1_TUPGB
16	57	23.4	90	1	GON3_DICLA
17	57	23.4	99	1	GON1_DICLA
18	57	23.4	379	1	METX_LEPME
19	55	22.5	249	1	PRA_MYCLE
20	55	22.5	266	1	XYNC_CALSA
21	55	22.5	480	1	Y523_XYLF
22	54	22.1	10	1	GON1_ALIMI
23	54	22.1	92	1	GON1_CHICK
24	53.5	21.9	148	1	HIL_MYTCA
25	53.5	21.9	202	1	HIL_MYTR
26	53.5	21.9	596	1	SDP_EIMBO
27	53.5	21.9	733	1	VINE_MOUSE
28	53.5	21.9	940	1	MAZ4_SCHCO
29	53	21.7	299	1	YKXC_CVAPA
30	53	21.7	485	1	RT16_MYXXA
31	53	21.7	575	1	ACEA_LYCES
32	52.5	21.5	1192	1	LMG2_ONCMA
33	52	21.3	82	1	GON3_ONCMA

RESULT 1

ID	TETX_CLOTE	STANDARD	PRT	1314 AA
AC	P04958;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).			
OS	Clostridium tetani.			
OG	Plasmid.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1513;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87053814; PubMed=3536478;			
RA	Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,			
RA	Weller U., Hudel M., Habermann E., Niemann H.;			
RT	"Tetanus toxin: primary structure, expression in E. coli, and			
RT	homology with botulinum toxins.";			
RL	EMBO J. 5:2495-2502(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CN3911;			
RX	MEDLINE=87040747; PubMed=3774547;			
RA	Fairweather N.F., Lyness V.A.;			
RT	"The complete nucleotide sequence of tetanus toxin.";			
RL	Nucleic Acids Res. 14:7809-7812(1986).			
RN	[3]			
RP	SEQUENCE OF 742-1314 FROM N.A.			
RX	MEDLINE=86085672; PubMed=3510187;			
RA	Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;			
RT	"Cloning, nucleotide sequencing, and expression of tetanus toxin			
RT	fragment C in Escherichia coli.";			
RL	J. Bacteriol. 165:21-27(1986).			
RN	[4]			
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.			
RX	MEDLINE=90201034; PubMed=2108021;			
RA	Kriegstein K., Henschen A., Weller U., Habermann E.;			
RT	"Arrangement of disulfide bridges and positions of sulfhydryl groups			
RT	in tetanus toxin.";			
RL	Eur. J. Biochem. 188:39-45(1990).			
RN	[5]			
RP	PARTIAL SEQUENCE.			
RX	MEDLINE=92037649; PubMed=1935979;			
RA	Kriegstein K.G., Henschen A.H., Weller U., Habermann E.;			
RT	"Limited proteolysis of tetanus toxin. Relation to activity and			
RT	identification of cleavage sites.";			
RL	Eur. J. Biochem. 202:41-51(1991).			
RN	[6]			
RP	IDENTIFICATION AS ZINC-PROTEASE.			
RX	MEDLINE=93010948; PubMed=1396558;			
RA	Schiavo G., Poullain B., Rossetto O., Benfenati F., Tauc L.,			
RA	Montecucco C.;			
RT	"Tetanus toxin is a zinc protein and its inhibition of			
RT	neurotransmitter release and protease activity depend on zinc.";			

34	52	21.3	89	1	GON3_PORNO	P51922 porichthys
35	52	21.3	90	1	GON3_PAGMA	P51921 pagrus major
36	52	21.3	90	1	GON3_SPAAU	P51923 sparus auratus
37	52	21.3	94	1	GON1_HAPBU	P51918 haplochromis
38	52	21.3	95	1	GON1_MORSA	P73812 morone saxatilis
39	52	21.3	95	1	GON1_PAGMA	P70074 pagrus major
40	52	21.3	95	1	GON1_SPAAU	P51919 sparus auratus
41	51.5	21.1	493	1	C6AD_DROME	O9449 drosophila
42	51.5	21.1	900	1	GUNH_CLOTH	P16218 clostridium
43	51	20.9	120	1	HV03_MOUSE	P01747 mus musculus
44	51	20.9	140	1	HV02_MOUSE	P01746 mus musculus
45	51	20.9	240	1	PRA_MYCTU	O53426 mycobacterium

ALIGNMENTS

RESULT 15

RHSHG

gonadoliberin - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
 C:Accession: A93780; A01411
 R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; Bl
 Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
 A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing facto
 A:Reference number: A93780; MUID:72094314
 A:Accession: A93780
 A:Molecule type: protein
 A:Residues: 1-10 <BUR>
 A:Note: the natural and synthetic hormones have the same biological activity
 C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.8%; Score 58; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 |||||
 Db 2 HWSYGLRPG 10

Search completed: October 10, 2002, 16:12:16
 Job time : 11.7906 secs

C;Accession: E95361
R.;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B.
. ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.
proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melli
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: E95361
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-374 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65455.1; PID:g14523923; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
peta, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium melliott.
A;Reference number: A96039; MUID:21368234; PMID:11474104
C;Contents: annotation
C;Genetics:
A;Gene: Smal1461
A;Genome: plasmid
C;Keywords: intramolecular lyase; isomerase

Query Match 24.0%; Score 58.5; DB 2; Length 374;
Best Local Similarity 41.2%; Pred. No.11;
Matches 14; Conservative 6; Mismatches 13; Indels 1; Gaps 1;

QY 2 HWSVGLRPGSGPSLQVIKANSKFIGITELSSGP 35
||||| | : : ||: ||| | :
Db 21 HWSYGIRE-SFVNLIEADGTVIGECTVAP 53

RESULT 14
RHPCG gonadoliblerin - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C;Accession: A01411
R;Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A;Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A;Reference number: A90172; MUID:72114303
A;Accession: A01411
A:Molecule type: protein
A;Residues: 1-10 <BAB>
R;Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A;Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A;Reference number: A90176; MUID:72065376
A;Contents: annotation; synthesis
A;Note: the synthetic and natural hormones have the same physicochemical and biologic
R;Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A;Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A;Reference number: A90175; MUID:72117544
A;Contents: annotation
A;Note: Trp-3 appears to be essential for biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 23.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSVGLRPG 10
|||||||
Db 2 HWSVGLRPG 10

[illegible]

GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 10.7137 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-19

Perfect score: 244

Sequence: 1 XHWSYGLRPGSGPSLQYIK.....GITELSGPSLHWSYGLRPX 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	31.1	1315	1 BTCLTN	tentoxylisin (EC 3.4.24.68) precursor - Clostridium tetani
2	69	28.3	82	2 I51365	gonadotropin-relea
3	65	26.6	74	2 I51092	gonadotropin relea
4	65	26.6	82	2 I51355	gonadotropin relea
5	65	26.6	82	2 I51331	gonadotropin relea
6	61	25.0	458	2 H87624	peptidase, M23/M37
7	60	24.6	447	2 B83563	conserved hypothet
8	59.5	24.4	66	2 S31029	gene 84 protein -
9	59.5	24.4	1172	2 T00065	hypothetical prote
10	59	24.2	123	2 G48677	Ig heavy chain V-D
11	59	24.2	135	2 PH1494	Ig heavy chain V r
12	58.5	24.0	90	1 RHMSG	gonadoliberin prec
13	58.5	24.0	374	2 E95361	probable muconate
14	58	23.8	10	1 RHPGG	gonadoliberin - pi
15	58	23.8	10	1 RHSHG	gonadoliberin - sh
16	58	23.8	67	2 I78541	gonadoliberin prec
17	58	23.8	89	2 I51423	gonadoliberin prec
18	58	23.8	92	1 RHHUG	gonadoliberin prec
19	58	23.8	92	1 RHRTG	gonadoliberin prec
20	57	23.4	112	2 C27887	Ig kappa chain V r
21	57	23.4	112	2 C27887	Ig kappa chain V r
22	57	23.4	115	2 S38715	Ig kappa chain V r
23	57	23.4	119	2 PH1518	Ig heavy chain V r
24	57	23.4	119	2 PH1519	Ig heavy chain V r
25	57	23.4	379	2 T44656	homoserine O-acety
26	57	23.4	530	2 F86467	hypothetical prote
27	56	23.0	102	2 PH1491	Ig heavy chain V r
28	56	23.0	120	2 A49043	Ig kappa chain V r
29	56	23.0	719	2 T52510	hypothetical prote

ALIGNMENTS

RESULT 1

BTCLTN

tentoxylisin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin

C:Species: Clostridium tetani

C:Date: 31-Mar-1988 #sequence.revision 31-Mar-1988 #text-change 18-Jun-1999

C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,

EMBO J. 5, 2495-2502, 1986

A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b

A:Reference number: A25689; MUID:87053814

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770

R:Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A:Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; MUID:87040747

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774

A:Experimental source: strain CN3911

R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.

J. Bacteriol. 165, 21-27, 1986

A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C i

A:Reference number: A25194; MUID:86085672

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <FA2>

A:Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921

A:Accession: B25194

A:Molecule type: protein

R:Matsumoto, M.; Lele, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.

Infect. Immun. 57, 3588-3593, 1989

A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termin

A:Reference number: A60759; MUID:90035436

A:Accession: A60759

A:Molecule type: protein

A:Residues: 461-475 <MAT>

R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.

J. Immunol. 142, 394-402, 1989

A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: JS0098; MUID:89093918

A:Contents: annotation; epitope region

R:Schiaivo, G.; Benfenati, F.; Poulin, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,

Nature 359, 832-835, 1992

A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo

A:Reference number: S27125; MUID:93063293

A:Contents: annotation

Db 1 FNNFTVSFWLRVPKVSASHLE 21

```

RESULT 15
US-09-382-855-5
; Sequence 5, Application US/09382855
; Patent No. 6174692
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie, Pierre;
; APPLICANT: Cerrottini, Jean-Charles; Carrel, Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor Rejection
; TITLE OF INVENTION: Antigen Precursors Mage-10, Antibodies Specific To The Molecule.
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/382,855
; FILING DATE: 25-August-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/089,595
; FILING DATE: 02-June-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Schofield, Mary Anne
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
US-09-382-855-5

```

Query Match 40.9%; Score 112; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 6.9e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 FNNFTVSFWLRVPKVSASHLE 37
 Db 1 FNNFTVSFWLRVPKVSASHLE 21

Search completed: October 10, 2002, 16:14:06
 Job time : 11.5769 secs

```

; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Golick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 243-0816
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-460-502-8

Query Match 40.9%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 13
US-08-724-774B-5
; Sequence 5, Application US/08724774B
; Patent No. 5908778
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
; APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
; APPLICANT: Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
; TITLE OF INVENTION: Rejection Antigen Precuros Mage-10,
; TITLE OF INVENTION: Antibodies Specific To The Molecule, and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felife & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6153728man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
; US-09-089-595-5

Query Match 40.9%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37

```

```

; NAME: Hanson, No. 5908778man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-724-774B-5

Query Match 40.9%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 14
US-09-089-595-5
; Sequence 5, Application US/09089595
; Patent No. 6153728
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coulie,
; APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
; APPLICANT: Stefan; Reed, Daryl
; TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
; TITLE OF INVENTION: Rejection Antigen Precuros Mage-10,
; TITLE OF INVENTION: Antibodies Specific To The Molecule, and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felife & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6153728man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
; US-09-089-595-5

Query Match 40.9%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPKVSASHLE 37

```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/668,381A
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,473
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,164
; REFERENCE/DOCKET NUMBER: 00786/269001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-668-381A-5

Query Match 41.6%; Score 114; DB 1; Length 618;
Best Local Similarity 95.5%; Pred. No. 2.4e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNTVSFWLVRPKVSASHLE 37
Db 249 MFNNTVSFWLVRPKVSASHLE 270

RESULT 10
US-07-610-525-1
; Sequence 1, Application US/07610525
; Patent No. 5196512
; GENERAL INFORMATION:
; APPLICANT: BIANCHI Eisabetta
; APPLICANT: PESSI Antonello
; APPLICANT: CORRADIN Giampietro
; TITLE OF INVENTION: SYNTHETIC PEPTIDES USEFUL AS
; TITLE OF INVENTION: UNIVERSAL CARRIERS FOR THE PREPARATION OF IMMUNOGENIC CONJUGATES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHEA & GOULD
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10020-1193
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/610,525
; FILING DATE: 19901108
; CLASSIFICATION: 424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-827-3000
; TELEFAX: 212-840-6702
; TELEX: 423973
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acid residues
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; FRAGMENT TYPE: internal fragment
```

```
US-07-610-525-1
Query Match 40.9%; Score 112; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLVRPKVSASHLE 37
Db 1 FNNFTVSFWLVRPKVSASHLE 21

RESULT 11
US-08-661-052-12
; Sequence 12, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-661-052-12

Query Match 40.9%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLVRPKVSASHLE 37
Db 1 FNNFTVSFWLVRPKVSASHLE 21

RESULT 12
US-08-460-502-8
; Sequence 8, Application US/08460502
; Patent No. 5843464
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
```

```
; Sequence 2, Application US/08280228
; Patent No. 5571694
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5571694th Glebe Road
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,228
; FILING DATE: 25-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/618,312
; FILING DATE: 27-NOV-1990
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mary J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-280-228-2

Query Match 41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 1.6e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LENFTVSWLVRPKVSASHLE 37
Db 83 MFNFTVSWLVRPKVSASHLE 104

RESULT 8
US-08-280-228-4
; Sequence 4, Application US/08280228
; Patent No. 5571694
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
```

```
; STREET: 1100 No. 5571694th Glebe Road
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,228
; FILING DATE: 25-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/618,312
; FILING DATE: 27-NOV-1990
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mary J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-280-228-4

Query Match 41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 1.6e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LENFTVSWLVRPKVSASHLE 37
Db 83 MFNFTVSWLVRPKVSASHLE 104

RESULT 9
US-08-668-381A-5
; Sequence 5, Application US/08668381A
; Patent No. 5780024
; GENERAL INFORMATION:
; APPLICANT: Brown, Robert H.
; APPLICANT: Fishman, Paul S.
; APPLICANT: Francis, Jonathan W.
; APPLICANT: Hosler, Betsy A.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-618-312A-2

Query Match 41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 1.6e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRPKVSASHLE 37
Db 83 MFNNFTVSFWLRPKVSASHLE 104

RESULT 5
US-07-618-312A-4
; Sequence 4, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
; US-07-618-312A-4

Query Match 41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 1.6e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRPKVSASHLE 37
Db 83 MFNNFTVSFWLRPKVSASHLE 104

RESULT 6
US-08-110-786A-8
; Sequence 8, Application US/08110786A
; Patent No. 5443966
; GENERAL INFORMATION:
; APPLICANT: FAIRWEATHER, Neil Fraser
; APPLICANT: MAKOFF, Andrew Joseph
; TITLE OF INVENTION: Expression of tetanus toxin fragment C
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 No. 5443966th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,786A
; FILING DATE: 23-AUG-1993 1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,337
; FILING DATE: 29-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00943
; FILING DATE: 20-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8914122.0
; FILING DATE: 20 June 1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary J. Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-110-786A-8

Query Match 41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 1.6e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRPKVSASHLE 37
Db 83 MFNNFTVSFWLRPKVSASHLE 104

RESULT 7
US-08-280-228-2
```

```
RESULT 2
US-08-488-351A-14
; Sequence 14, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Y1
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-14
Query Match 55.5%; Score 152; DB 2; Length 32;
Best Local Similarity 87.9%; Pred. No. 4.4e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 17 FNNFTVFWLRVPKVSASHLEGPLSHWSYGLRP 49
| | | | | | | | | | | | | | | | | | | | | |
Db 3 FNNFTVFWLRVPKVSASHLE---HWSYGLRP 31

RESULT 3
PCT-US93-11703-64
; Sequence 64, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
```

```
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-11703-64
Query Match 41.6%; Score 114; DB 5; Length 31;
Best Local Similarity 95.5%; Pred. No. 6e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVFWLRVPKVSASHLE 37
| | | | | | | | | | | | | | | | | |
Db 6 MFNNFTVFWLRVPKVSASHLE 27

RESULT 4
US-07-618-312A-2
; Sequence 2, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 10.5769 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-18
Perfect score: 274
Sequence: 1 XHWSYGLRPGSGPSLFNPF.....VSASHLEGPSLHWSYGLRFX 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	55.5	32	1	US-08-446-692-14
2	152	55.5	32	2	US-08-488-351A-14
3	114	41.6	31	5	PCT-US93-11703-64
4	114	41.6	452	1	US-07-618-312A-2
5	114	41.6	452	1	US-07-618-312A-4
6	114	41.6	452	1	US-08-110-786A-8
7	114	41.6	452	1	US-08-280-228-2
8	114	41.6	452	1	US-08-280-228-4
9	114	41.6	618	1	US-08-668-381A-5
10	112	40.9	21	1	US-07-610-525-1
11	112	40.9	21	2	US-08-661-052-12
12	112	40.9	21	2	US-08-460-502-8
13	112	40.9	21	2	US-08-724-774B-5
14	112	40.9	21	4	US-09-089-595-5
15	112	40.9	21	4	US-09-382-855-5
16	112	40.9	21	4	US-09-183-714B-5
17	112	40.9	21	4	US-09-188-082-12
18	112	40.9	21	4	US-09-171-969-10
19	112	40.9	21	5	PCT-US93-11703-66
20	107	39.1	22	2	US-08-446-692-5
21	107	39.1	22	2	US-08-488-351A-5
22	107	39.1	22	3	US-09-100-409A-41
23	107	39.1	22	5	PCT-US95-13841-8
24	102.5	37.4	699	2	US-08-694-865-16
25	102.5	37.4	699	3	US-09-124-491-16
26	101	36.9	49	1	US-08-387-156-4
27	101	36.9	49	2	US-08-694-865-4

28 101 36.9 49 2 US-08-878-748-4 Sequence 4, Appli
29 101 36.9 49 3 US-09-124-491-4 Sequence 4, Appli
30 101 36.9 544 1 US-08-387-156-10 Sequence 10, Appl
31 101 36.9 544 2 US-08-694-865-10 Sequence 10, Appl
32 101 36.9 544 2 US-08-878-748-10 Sequence 10, Appl
33 101 36.9 544 3 US-09-124-491-10 Sequence 10, Appl
34 101 36.9 977 1 US-08-387-156-8 Sequence 8, Appli
35 101 36.9 977 2 US-08-694-865-8 Sequence 8, Appli
36 101 36.9 977 3 US-09-124-491-8 Sequence 8, Appli
37 101 36.9 977 1 US-07-610-525-2 Sequence 2, Appli
38 100 36.5 19 4 US-09-026-276-35 Sequence 35, Appl
39 97 35.4 40 4 US-09-026-276-34 Sequence 34, Appl
40 97 35.4 41 4 US-07-690-983D-45 Sequence 45, Appl
41 96 35.0 44 1 US-07-690-983D-47 Sequence 47, Appl
42 96 35.0 84 1 US-08-458-814-6 Sequence 6, Appli
43 95 34.7 52 3 US-08-458-814-7 Sequence 7, Appli
44 95 34.7 55 3 US-08-319-704-11 Sequence 11, Appl
45 94.5 34.5 20 2

ALIGNMENTS

RESULT 1
US-08-446-692-14
; Sequence 14, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF INVENTIONS: and synthetic universal immune stimulants for vaccines
; NUMBER OF SEQUENCE: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-14

Query Match 55.5%; Score 152; DB 1; Length 32;
Best Local Similarity 87.9%; Pred. No. 4.4e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 17 FNNFTVSFWLRPKVSASHLEGPSLHWSYGLRFP 49
| | | | | | | | | | | | | | | | | | | | | |
Db 3 FNNFTVSFWLRPKVSASHLE---HWSYGLRFP 31


```
CC B-cell group derived from the cell-associated PA; and (2) at least 1
CC first T helper cell group which is foreign to the animal. Analogues of
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC part of all known and predicted CTL and B-cell epitopes of the respective
CC PA and including at least one foreign T helper epitope are also claimed.
CC The method is used to treat prostate, prostate/breast or breast cancer
CC when the PA is human PSM, FGF8b and Her2, respectively.
CC Note: This sequence was constructed from the wild type human PSM
CC (RAY92619), which appears on pages 184-187 of the specification.
XX
SQ Sequence 693 AA;
Query Match 43.2%; Score 118.5; DB 21; Length 693;
Best Local Similarity 66.7%; Pred. No. 3.6e-07;
Matches 24; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Qy 13 GPSLFNNFTVSFWLRVPKVSASHLEGGPSLHWSYGLR 48
| | ||||| ||||| ||||| ||||| : : : : :
Db 149 GKVFENFTVSFWLRVPKVSASHLE-PADYFAPGVK 183

Search completed: October 10, 2002, 16:05:16
Job time : 27.2821 secs
```

CC Her2, respectively.
 XX SQ Sequence 31 AA; Query Match 43.2%; Score 118.5; DB 21; Length 31;
 Best Local Similarity 92.3%; Pred. No. 9.4e-09; Matches 24; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 17 FNNFTVSFWLRVFKVSASHLEGPSLH 42
 Db 6 FNNFTVSFWLRVFKVSASHLE-PSSH 30

RESULT 13
 AAW81332
 ID AAW81332 standard; Protein; 158 AA.
 XX AC AAW81332;
 XX DT 21-APR-1999 (first entry)
 DE TNF30-1, a TNF-alpha analogue.
 XX Human tumour necrosis factor-alpha; TNF-alpha: TNF-alpha analogue;
 KW vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
 KW cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
 KW asthma.
 XX Synthetic.
 OS Homo sapiens.
 OS WO9846642-A1.
 XX 22-OCT-1998.
 XX 15-APR-1998; 98WO-DK00157.
 XX 24-APR-1997; 97US-0044187.
 PR 15-APR-1997; 97DK-0000418.
 XX (PERR) FARM LAB FERRING AS.
 PA Dalum I, Elsner H, Jensen MR, Mouritsen S;
 PI WPI; 1998-594561/50.
 DR N-PSDB; AAV68421.
 XX Modified human tumour necrosis factor-alpha - comprises
 PT immunodominant T cell epitope, useful in vaccines to treat or
 PT prevent TNF-associated diseases, e.g. cancer
 XX Example 1; Page 74-75; 134pp; English.
 PS The present sequence represents a modified human tumour necrosis
 CC factor-alpha (TNF-alpha) analogue. The analogues have no residual
 CC TNF activity and are immunogenic in a large proportion of the human
 CC population (by using promiscuous epitopes). The TNF-alpha analogue
 CC is able to generate, in humans, neutralizing antibodies to wild-type
 CC human TNF alpha, has at least one fragment of TNF substituted by a
 CC peptide containing an immunodominant T-cell epitope, and at least one
 CC TNF-alpha B-cell epitope. The substitution causes a significant change
 CC in the amino acid sequence of any one of the strands in the front
 CC beta-sheet, any of the connecting loops or any of the B', I or D strands
 CC in the back beta-sheet. The TNF-alpha analogues are used as vaccines for
 CC treatment or prevention of diseases associated with excessive release
 CC or activity of TNF-alpha, e.g. rheumatoid arthritis, Crohn's disease,
 CC ulcerative colitis, cancer of any sort, disseminated sclerosis, diabetes,
 CC psoriasis, osteoporosis and asthma.
 XX SQ Sequence 158 AA;
 Query Match 43.2%; Score 118.5; DB 19; Length 158;
 Best Local Similarity 80.6%; Pred. No. 6.3e-08;

Matches 25; Conservative 1; Mismatches 4; Indels 1; Gaps 1;
 QY 7 LRPGSSGSPSLFNNFTVSFWLRVFKVSASHLE 37
 Db 2 VRSSSRTPS-FNNFTVSFWLRVFKVSASHLE 31

RESULT 14
 AAY92647
 ID AAY92647 standard; Protein; 693 AA.
 XX AC AAY92647;
 XX DT 10-AUG-2000 (first entry)
 XX Mutant human PSM antigen splice variant construct, hPSM/6.3.
 KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer; PSM;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX Homo sapiens.
 OS Synthetic.
 OS Key Location/Qualifiers
 PH Peptide 153..173
 FT /label= P30
 FT /note= "foreign epitope"
 FT 391..405
 FT /label= P2
 FT /note= "foreign epitope"
 XX WO200020027-A2.
 XX 13-APR-2000.
 XX 05-OCT-1999; 99WO-DK00525.
 XX 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX (MEBI-) M & E BIOTECH AS.
 XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX WPI; 2000-349917/30.
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX Example 1; Page -: 220pp; English.
 PS AAY92627-49 are mutant immunogenized human prostate specific membrane
 CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
 CC P30). The immunogenic analogues of PSM can be used in the claimed method
 CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
 CC antibody binding regions and cysteine residues involved in disulfide
 CC bonds are preserved in the immunogenized forms. The method is used for
 CC inducing immune responses against weakly immunogenic cell-associated
 CC peptide antigens (PA) such as those associated with cancers
 CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
 CC method comprises effecting simultaneous presentation by antigen producing
 CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
 CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the respective
 CC PA and including at least one foreign T helper epitope are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer
 CC when the PA is human PSM, FGF8b and Her2, respectively.

PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -

XX
XX Example 7; Page 134; 172pp; English.

XX The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.

XX
XX SQ Sequence 122 AA;

Query Match 43.4%; Score 119; DB 21; Length 122;
Best Local Similarity 81.5%; Pred. No. 4e-08; Mismatches 2; Indels 0; Gaps 0;

XX
XX Matches 22; Conservative 3;

QY 11 SSGPSLFNFTVFWLRPKVSASHLE 37
:|:|||||||||||||||||
Db 24 TSNETMENFTVFWLRPKVSASHLE 50

RESULT 11
AAB45524
ID AAB45524 standard; Protein; 122 AA.

XX
XX AAB45524;

XX
XX 26-FEB-2001 (first entry)

XX
XX Modified murine interleukin-5 SEQ ID NO: 48.

XX
XX Asthma; IL-5; interleukin-5; allergy; cytokine; helminthic infection;
XX cancer; eosinophilia; vaccine; allergic rhinitis.

XX
XX Mus musculus.

OS
OS Clostridium tetani.

XX
XX WO200065058-A1.

XX
XX 02-NOV-2000.

XX
XX 19-APR-2000; 2000WO-DK00205.

XX
XX 23-APR-1999; 99DK-0000552.

PR
PR 06-MAY-1999; 99US-0132811.

XX
XX (MEBI-) M & E BIOTECH AS.

XX
XX Klysner S;

XX
XX WPI; 2000-672791/65.

DR
DR N-PSDB; AAC68877.

XX
XX Down-regulating interleukin 5 (IL-5) activity in humans by
PT administering IL-5 and/or an IL-5 analogue, useful in the treatment,
PT prophylaxis or amelioration of asthma or other chronic allergic
PT conditions -

XX
XX Example 7; Page 156; 172pp; English.

XX The present invention is concerned with methods of treating asthma,
CC eosinophilia, allergic rhinitis and other allergic diseases. These
CC involve the use of interleukin-5 (IL-5) analogues and modified IL-5
CC proteins and their coding sequences to down-regulate IL-5 activity and
CC thus reduce eosinophil numbers. The allergic diseases may be treated
CC using autovaccines, nucleic acid vaccines or live vaccines. In addition,
CC it is possible that they may be used in the treatment of cancer and
CC helminthic infections.

XX
XX SQ Sequence 122 AA;

Query Match 43.4%; Score 119; DB 21; Length 122;
Best Local Similarity 81.5%; Pred. No. 4e-08; Mismatches 2; Indels 0; Gaps 0;

XX
XX Matches 22; Conservative 3;

QY 11 SSGPSLFNFTVFWLRPKVSASHLE 37
:|:|||||||||||||||||
Db 24 TSNETMENFTVFWLRPKVSASHLE 50

RESULT 12
AAY92655
ID AAY92655 standard; Peptide; 31 AA.

XX
XX AAY92655;

XX
XX 10-AUG-2000 (first entry)

XX
XX PSMpep012 - P30 inserted in hPSM insertion position 10.

XX
XX Foreign epitope; P2: prostate specific membrane antigen; vaccination;
KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
KW prostate cancer; cell-associated peptide antigen.

XX
XX Synthetic.

OS
OS Homo sapiens.

XX
XX Key Location/Qualifiers

PH Peptide 6..26
FT /label= P30

XX
XX WO200020027-A2.

XX
XX 13-APR-2000.

XX
XX 05-OCT-1999; 99WO-DK00525.

XX
XX 05-OCT-1998; 98DK-0001361.

PR
PR 20-OCT-1998; 98US-0105011.

XX
XX (MEBI-) M & E BIOTECH AS.

XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;

XX
XX WPI; 2000-349917/30.

XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer

XX
XX Example 1; Page 118; 220pp; English.

XX
XX AAY92650-55 are peptides designed which correspond to the P2 and P30
CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
CC amino acids in each end. The flanking amino acids correspond to the
CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g.
CC T cell proliferation assays, but also for ELISA or other in vitro
CC assays. The claims detail a method for inducing immune responses against
CC weakly immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, hPSM,
CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
CC method comprises effecting simultaneous presentation by antigen
CC producing cells (APCs) of the animals immune system of: (1) at least 1
CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC B-cell group derived from the cell-associated PA; and (2) at least 1
CC first T helper cell group which is foreign to the animal. Analogues of
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC part of all known and predicted CTL and B-cell epitopes of the
CC respective PA and including at least one foreign T helper epitope (e.g.
CC P2 and/or P30) are also claimed. The method is used to treat prostate,
CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and


```

XX AAY92627;
XX 10-AUG-2000 (first entry)
XX Mutant human prostate specific membrane antigen construct, hPSM1.1.
DE
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Peptide 17..31
FT /label= P2
FT /note= "foreign epitope"
FT
FT Peptide 32..52
FT /label= P30
FT /note= "foreign epitope"
FT
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
PA
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
CC P30). The immunogenic analogues of PSM can be used in the claimed method
CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
CC antibody binding regions and cysteine residues involved in disulfide
CC bonds are preserved in the immunogenized forms. The method is used for
CC inducing immune responses against weakly immunogenic cell-associated
CC peptide antigens (PA) such as those associated with cancers
CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
CC method comprises effecting simultaneous presentation by antigen producing
CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC B-cell group derived from the cell-associated PA; and (2) at least 1
CC first T helper cell group which is foreign to the animal. Analogues of
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC part of all known and predicted CTL and B-cell epitopes of the respective
CC PA and including at least one foreign T helper epitope are also claimed.
CC The method is used to treat prostate, prostate/breast or breast cancer
CC when the PA is human PSM, FGF8b and Her2, respectively.
CC Note: This sequence was constructed from the wild type human PSM
CC (AAY92619), which appears on pages 184-187 of the specification.
XX
XX Sequence 750 AA;
XX
XX Query Match 44.2%; Score 121; DB 21; Length 750;
XX Best Local Similarity 95.8%; Pred. No. 1.8e-07;
XX Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 16 LFNNFTVSEWLRVPKVSASHLEGP 39

```

```

DB 31 LFNNFTVSEWLRVPKVSASHLETP 54
|||||
RESULT 6
AAB20149
ID AAB20149 standard; Protein; 109 AA.
XX
XX AAB20149;
XX
XX 30-APR-2001 (first entry)
XX
XX Growth differentiation factor 8 AutoVac construct GDF-8 P30-2.
DE
XX
XX Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX
XX Chimeric - Homo sapiens.
OS Chimeric - Clostridium tetani.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..48
FT /note= "identical to residues 267-314 of human
FT GDF-8"
FT
FT Region 49..69
FT /note= "tetanus toxoid P2 epitope"
FT Region 70..109
FT /note= "identical to residues 336-375 of human
FT GDF-8"
FT
FT Misc-difference 73
FT /note= "Cys-73 may be substituted by Ser to avoid
FT disulfide bond formation"
FT Misc-difference 90..91
FT /note= "optionally replaced by Glu-Gly"
FT
XX WO200105820-A2.
XX
XX 25-JAN-2001.
XX
XX 20-JUL-2000; 2000WO-DK00413.
XX
XX 20-JUL-1999; 99DK-0001014.
PR 26-JUL-1999; 99US-0145275.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Halkier T, Mouritsen S, Klysner S;
PI
PI WPI; 2001-112680/12.
DR
XX
XX Increasing the muscle mass of animals used in meat production by down
PT regulating growth differentiation factor 8 (GDF-8) activity in the
PT animal through induction of anti-GDF-8 antibody production -
XX
XX Example 1; Page 101-102; 110pp; English.
XX
XX The present sequence is that of AutoVac construct GDF-8 P30-2,
CC comprising the 109 C-terminal amino acid residues of human
CC growth differentiation factor 8 (GDF-8) in which residues 49-69 are
CC replaced by the promiscuous tetanus toxin T-cell epitope P30 (see
CC AAB20144). It is an object of the invention to produce a
CC recombinant therapeutic vaccine that is capable of effecting
CC down-regulation of GDF-8 in order to increase the muscle growth
CC rate of farm animals. The vaccines (see AAB20145-53) are capable
CC of breaking autotolerance against autologous GDF-8. They comprise
CC the C-terminal portion of human GDF-8 in which a portion of the
CC native sequence is replaced by a T-cell epitope such as P30, with
CC minimal disturbance of the authentic 3-dimensional structure of
CC the protein. Nucleic acids encoding the GDF-8 variants can be used
CC for genetic immunisation of the animals. Down-regulation of GDF-8
CC activity can increase muscle mass by up to at least 45% in cattle,
CC

```

```

FT Peptide, 24..38
FT /label= p2
FT /note= "foreign epitope"
FT Peptide
FT 673..693
FT /label= p30
FT /note= "foreign epitope"
XX
PN WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
CC P30). The immunogenic analogues of PSM can be used in the claimed method
CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
CC antibody binding regions and cysteine residues involved in disulfide
CC bonds are preserved in the immunogenized forms. The method is used for
CC inducing immune responses against weakly immunogenic cell-associated
CC peptide antigens (PA) such as those associated with cancers
CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
CC method comprises effecting simultaneous presentation by antigen producing
CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC B-cell helper cell group which is cell-associated PA; and (2) at least 1
CC first T helper cell group which is foreign to the animal. Analogues of
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC part of all known and predicted CTL and B-cell epitopes of the respective
CC PA and including at least one foreign T helper epitope are also claimed.
CC The method is used to treat prostate, prostate/breast or breast cancer
CC when the PA is human PSM, FGF8b and Her2, respectively.
CC Note: This sequence was constructed from the wild type human PSM
CC (AAY92619), which appears on pages 184-187 of the specification.
XX
XX Sequence 750 AA;
XX
XX Query Match 44.3%; Score 121.5; DB 21; Length 750;
XX Best Local Similarity 86.2%; Pred. No. 1.6e-07;
XX Matches 25; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
XX
XX QY 17 FNNFTVFWLRVPKVSASHLEGPSLHWSY 45
XX |||||
XX Db 673 FNNFTVFWLRVPKVSASHLE-PSSHNY 700
XX
XX RESULT 4
XX AAY92665
XX ID AAY92665 standard; Peptide; 216 AA.
XX
XX AC AAY92665;
XX
XX DT 10-AUG-2000 (first entry)
XX
XX MUC-1 analogue containing foreign epitopes.
XX
XX DE Mucin repeat; MUC-1 analogue; vaccination; self-protein; cancer;
XX
XX KW

```

```

KW cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;
KW cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 61..75
XX /label= p2
XX Peptide 136..156
XX /label= p30
XX /note= "q"
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 4; Page -: 220pp; English.
XX
XX This is an immunogenized MUC-1 analogue containing foreign epitopes P2
CC and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate
CC specific membrane antigen (hPSM) can be used in the claimed method as an
CC autovaccine to induce a CTL response. Subdominant CTL epitopes,
CC antibody binding regions and cysteine residues involved in disulfide
CC bonds are preserved in the immunogenized forms (see features table). 10
CC regions suitable for the insertion of foreign T helper epitopes were
CC identified. The method is used for inducing immune responses against
CC weakly immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)
CC and/or fibroblast growth factor 8b (FGF8b). The method comprises
CC effecting simultaneous presentation by antigen producing cells (APCs) of
CC the animals immune system of: (1) at least 1 CTL (cytotoxic
CC T-lymphocyte) group derived from the PA and/or at least 1 B-cell helper
CC derived from the cell-associated PA; and (2) at least 1 first T helper
CC cell group which is foreign to the animal. Analogues of human PSM, human
CC Her2 and human/murine FGF8b comprising a substantial part of all known
CC and predicted CTL and B-cell epitopes of the respective PA and including
CC at least one foreign T helper epitope are also claimed. The method is
CC used to treat prostate, prostate/breast or breast cancer when the PA is
CC human PSM, FGF8b and Her2, respectively.
CC Note: This sequence does not appear in the specification. It was made
CC using the mucin repeat sequence (AAY92664), P2 and P30 (AAY92625-26),
CC which appear on pages 220, 213 and 214 respectively, of the
XX specification.
XX
XX Sequence 216 AA;
XX
XX Query Match 44.2%; Score 121; DB 21; Length 216;
XX Best Local Similarity 64.1%; Pred. No. 4.2e-08;
XX Matches 25; Conservative 2; Mismatches 2; Indels 10; Gaps 1;
XX
XX QY 9 PGSSGP-----SLFNNFTVFWLRVPKVSASHLE 37
XX |||||
XX Db 118 PGSTAPPANGVTSAPDTFRNFTVFWLRVPKVSASHLE 156
XX
XX RESULT 5
XX AAY92627
XX ID AAY92627 standard; Protein; 750 AA.

```


FT	XX		/note= "optionally replaced by Glu-Gly"
PN	XX	WO200105820-A2.	
PD	XX	25-JAN-2001.	
PF	XX	20-JUL-2000; 2000WO-DK00413.	
PR	XX	20-JUL-1999; 99DK-0001014.	
PR	XX	26-JUL-1999; 99US-0145275.	
PA	XX	(MEBI-) M & E BIOTECH AS.	
PI	XX	Halkier T, Mouritsen S, Klysner S;	
WI	XX	WPI; 2001-112680/12.	
XX	XX		
PT	PT	Increasing the muscle mass of animals used in meat production by down-	
PT	PT	regulating growth differentiation factor 8 (GDF-8) activity in the	
PT	PT	animal through induction of anti-GDF-8 antibody production -	
XX	XX		
PS	XX	Example 1; Page 102-103; 110pp; English.	
XX	XX		
CC	CC	The present sequence is that of AutoVac construct GDF-8 P30-3A,	
CC	CC	comprising the 109 C-terminal amino acid residues of human	
CC	CC	growth differentiation factor 8 (GDF-8) in which residues 79-99 are	
CC	CC	replaced by the promiscuous tetanus toxin T-cell epitope P30 (see	
CC	CC	AAB20144). It is an object of the invention to produce a	
CC	CC	recombinant therapeutic vaccine that is capable of effecting	
CC	CC	down-regulation of GDF-8 in order to increase the muscle growth	
CC	CC	rate of farm animals. The vaccines (see AAB20145-53) are capable	
CC	CC	of breaking autotolerance against autologous GDF-8. They comprise	
CC	CC	the C-terminal portion of human GDF-8 in which a portion of the	
CC	CC	native sequence is replaced by a T-cell epitope such as P30, with	
CC	CC	minimal disturbance of the authentic 3-dimensional structure of	
CC	CC	the protein. Nucleic acids encoding the GDF-8 variants can be used	
CC	CC	for genetic immunisation of the animals. Down-regulation of GDF-8	
CC	CC	activity can increase muscle mass by up to at least 45% in cattle,	
CC	CC	pigs and poultry used for meat production, reducing the need for	
CC	CC	antibiotic feed-additives. Anti-GDF8 vaccines can be used to	
CC	CC	treat human diseases such as cancer cachexia where muscle atrophy is	
CC	CC	pronounced and for patients suffering from acute and chronic heart	
CC	CC	failure.	
XX	XX		
SQ	XX	Sequence 109 AA;	
		Query Match 44.9%; Score 123; DB 22; Length 109;	
		Best Local Similarity 78.1%; Pred. No. 1e-08;	
		Matches 25; Conservative 2; Mismatches 1; Indels 4; Gaps	
QY	10	GSSG----PSLFNNFTVSEFWLVPKVSASHLE 37	
		:	
Db	68	GSAGPCCTPTKFNFTVSEFWLVPKVSASHLE 99	
RESULT 3			
AAI92633		RAY92633	
ID	XX	RAY92633 standard; Protein; 750 AA.	
AC	XX	RAY92633;	
XX	XX		
DT	XX	10-AUG-2000 (first entry)	
DE	XX	Mutant human prostate specific membrane antigen construct, hPSM1.10.	
KW	XX	Prostate specific membrane antigen; immunogenized construct; mutant;	
KW	XX	vaccination; cytotoxic T-lymphocyte immunity; breast cancer;	
KW	XX	prostate cancer; cell-associated peptide antigen; foreign epitope.	
OS	XX	Homo sapiens.	
OS	XX	Synthetic.	
XX	XX		
Key	XX	Location/Qualifiers	


```
RESULT 2
US-08-488-351A-27
; Sequence 27, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang YI
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-27

Query Match 48.1%; Score 127.5; DB 2; Length 33;
Best Local Similarity 81.8%; Pred. No. 6.6e-10;
Matches 27; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 18 EKKIAKMEKASSVFNVNSSGSLHWSYGLRP 50
DB 3 EKKIAKMEKASSVFNVNSSGGE--HWSYGLRP 32
|||||
|||||

RESULT 3
US-08-313-288B-18
; Sequence 18, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-18

Query Match 37.4%; Score 99; DB 1; Length 412;
Best Local Similarity 56.1%; Pred. No. 6.9e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSLD-----EKKIAKMEKASSVFNVNSSG 39
DB 361 IKPGSANKPKDELVDYENDIEKKICKMERKCCSVFNVNSSIG 401
|||||
|||||

RESULT 4
US-08-760-797A-3
; Sequence 3, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HbsAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 10.7885 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-20

Perfect score: 265

Sequence: 1 XHWSYGLRPGSGPSLDEKK.....NVNSSGSPSLHWSYGLRXP 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127.5	48.1	33	1	US-08-446-692-27
2	127.5	48.1	33	2	US-08-488-351A-27
3	99	37.4	412	1	US-08-313-288B-18
4	99	37.4	424	2	US-08-760-797A-3
5	99	37.4	424	4	US-08-932-929B-3
6	97.5	36.8	49	1	US-08-387-156-4
7	97.5	36.8	49	2	US-08-694-865-4
8	97.5	36.8	49	2	US-08-878-748-4
9	97.5	36.8	49	3	US-09-124-491-4
10	97.5	36.8	544	1	US-08-387-156-10
11	97.5	36.8	544	2	US-08-694-865-10
12	97.5	36.8	544	2	US-08-878-748-10
13	97.5	36.8	544	3	US-09-124-491-10
14	97.5	36.8	699	2	US-08-694-865-16
15	97.5	36.8	699	2	US-08-932-929B-16
16	97.5	36.8	977	1	US-08-387-156-8
17	97.5	36.8	977	2	US-08-694-865-8
18	97.5	36.8	977	2	US-08-878-748-8
19	97.5	36.8	977	2	US-09-124-491-8
20	91	34.3	423	2	US-08-760-797A-1
21	91	34.3	424	4	US-08-932-929B-1
22	89	33.6	21	1	US-08-186-266-6
23	89	33.6	21	1	US-08-446-692-48
24	89	33.6	21	2	US-08-488-351A-48
25	89	33.6	21	3	US-09-100-409A-54
26	89	33.6	21	5	PCT-US95-02121-97
27	89	33.6	21	5	PCT-US95-13841-20

28	88.5	33.4	40	4	US-09-026-276-35	Sequence 35, Appl
29	88.5	33.4	41	4	US-09-026-276-34	Sequence 34, Appl
30	87.5	33.0	44	1	US-07-690-983D-45	Sequence 45, Appl
31	87.5	33.0	52	3	US-08-458-814-6	Sequence 6, Appl
32	87.5	33.0	55	3	US-08-458-814-7	Sequence 7, Appl
33	87.5	33.0	84	1	US-07-690-983D-47	Sequence 47, Appl
34	84.5	31.9	20	1	US-07-690-983D-40	Sequence 40, Appl
35	84.5	31.9	20	4	US-09-026-276-26	Sequence 26, Appl
36	84.5	31.9	20	4	US-09-026-276-29	Sequence 29, Appl
37	84.5	31.9	20	4	US-09-026-276-30	Sequence 30, Appl
38	84.5	31.9	20	4	US-09-026-276-31	Sequence 31, Appl
39	84.5	31.9	24	1	US-07-690-983D-43	Sequence 43, Appl
40	79	29.8	17	5	PCT-US95-02121-96	Sequence 96, Appl
41	75	28.3	16	2	US-08-817-933A-7	Sequence 7, Appl
42	73.5	27.7	20	1	US-08-465-167A-20	Sequence 20, Appl
43	73.5	27.7	20	5	PCT-US92-07218-17	Sequence 17, Appl
44	71	26.8	17	1	US-08-188-223-6	Sequence 6, Appl
45	71	26.8	17	4	US-08-968-466-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-08-446-692-27
; Sequence 27, Application US/08446692
; Patent No. 5759551

; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna

; APPLICANT: Wang, Chang Yi

; APPLICANT: Zamb, Timothy

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin

; REGISTRATION NUMBER: 29,323

; REFERENCE/DOCKET NUMBER: 1151-4146 US2

; TELEPHONE: (212)415-8745

; TELEFAX: (516)751-6849

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 33 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-446-692-27

Query Match 48.1%; Score 127.5; DB 1; Length 33;
Best Local Similarity 81.8%; Pred. No. 6.6e-10;
Matches 27; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 18 EKKIAMKASSVNVNSSGSPSLHWSYGLR 50

DB 3 EKKIAMKASSVNVNSSGSPSLHWSYGLR 32

AA58135 1
ID AA58135 standard; Protein; 49 AA.
XX
AC AA58135;
XX
DT 07-MAR-2000 (first entry)
XX
DE GnRH analogue multimer, containing four copies of the GnRH analogue.
XX
KW Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
KW non-androgenic; steroid; reduction; weight gain; muscle distribution;
KW fat distribution; male pattern; boar taint; flavour; impairment;
KW reliable; immunocastration; meat production.
XX
OS Synthetic.
OS Mammalia.
XX
PN W09956771-A2.
XX
PD 11-NOV-1999.
XX
PF 05-MAY-1999; 99WO-CA00360.
XX
PR 05-MAY-1998; 98US-0084217.
XX
PA (BIOS-) BIOTAR INC.
XX
PI Manns JG, Acres SD, Harland R;
XX
DR WPI; 2000-062125/05.
XX
DR N-PSDB; AA46402.
XX
PT Production of uncastrated male food animals using vaccines -
XX
PS Example 1; Fig 2b; 87pp; English.
XX
CC This sequence represents four copies of a gonadotropin
CC releasing hormone (GnRH) analogue, DNA encoding which was
CC used in the construction of a chimeric GnRH-leukotoxin (LKT)
CC fusion gene (AA46400). This fusion gene encodes a GnRH-LKT fusion
CC protein which may be used as a vaccine. The LKT portion of the protein
CC acts to enhance the immunogenicity of the GnRH portion. The invention
CC relates to a method of using two GnRH immunogen vaccines to produce
CC uncastrated male animals for meat production, one vaccination prior to
CC or during the fattening period to reduce circulating testosterone
CC levels, and the second vaccination about 2-8 weeks before slaughter to
CC substantially reduce androgenic and/or non-androgenic steroids. The
CC invention is used to produce food animals that exhibit the weight gain
CC and muscle/fat distribution of male animals without the problems
CC associated with male animals. Such problems include "boar taint", a
CC urine-like odour found in cooked meat of uncastrated pigs which is
CC caused by steroids stored in the tissues, and similar flavour
CC impairments in the meat of other intact male animals. The invention is
CC more reliable than prior art immunocastration techniques.
XX
SQ Sequence 49 AA;
Query Match 36.8%; Score 97.5; DB 21; Length 49;
Best Local Similarity 42.9%; Pred. No. 1.7e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;
QY 2 HWSYGLRFGSGPSLDEKKAEMKASSFNVVNSSGSPSLHWSYGLRP 50
|||||||
DB 2 HWSYGLRFGS-----GSDWSYGLRFGGSSQHWYGLRP 35

Db 2 HWSYGLRPGS-----GSODWSYGLRPGSSOHWSYGLRP 35

RESULT 13

AAW61542
ID AAW61542 standard; Protein: 49 AA.

AC AAW61542;

DT 27-OCT-1998 (first entry)

DE Peptide hormone GnRH-2 decapeptide (4 copies) fragment.

GnRH; gonadotrophin releasing hormone peptide hormone; leukotoxin;
 KW
 immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;
 KW
 immune response; hormone receptor; cancerous cell; domestic animal;
 KW
 porcine; bovine; luteinizing hormone; follicle stimulating hormone;
 KW
 immunocastrate.

OS Synthetic.

PN WO9834639-A1.

13-AUG-1998

AA
PF 04-FEB-1998: 98WO-CA00059.

05-FEB-1997: 97US-0036883.

PA (BIOS-) BIOSTAR INC.

PI Acres SD, Harland R, Manns JG;

WPI: 1998-446952/38.

DR N-PSDB; AAV45190.

PT Immunisation against endogenous molecules by administering vaccine
PT to ear - useful to elicit efficient and uniform immune response
PT against e.g. gonadotrophin releasing hormone to immunocastrate pigs
PT and cattle

Example 1; Fig 1B; 61pp; English.

This represents the amino acid sequence of the gonadotrophin releasing hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-GnRH polypeptide gene fusions. This is used to exemplify the method of invention of immunisation against endogenous molecules by administering a vaccine which comprises an immunogen and a carrier to the ear of the mammal. The method is useful for eliciting an efficient and uniform immune response to block or suppress the activity of an endogenous hormone, hormone receptor, agonist or antagonist in a vaccinated subject, or to elicit an immune response against a targeted endogenous cell type (e.g. a cancerous or otherwise diseased cell). It is especially useful to reduce the levels of GnRH in domestic animals, especially in porcine or bovine species. The use of GnRH immunogens in the vaccine reduces the levels of luteinizing hormone and follicle stimulating hormone and helps in immunocastrating the animal. Administration of vaccine compositions to the ear instead of intramuscular administration into the neck increases the efficiency of vaccination of mammals against endogenous immunogens, and may increase uniformity of vaccine presentation since the ear is relatively uniform from animal to animal.

Sequence 49 AA:

Query Match 36.8%; Score 97.5; DB 19; Length 49;
Best Local Similarity 42.9%; Pred. No. 1.7e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 1

Qy 2 HWSYGLRPGSSGSLDEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
 *||| ||||| | :| | | | | | |
Db 2 HWSYGLRPGS-----GSODWSYGLRPGSGSSHWSYGLRP 35

RESULT 14

RESOLUTION 14
 AAY58363
 ID AAY58363 standard; Protein; 49 AA.

AC AAY58363;

DT 27-MAR-2000 (first entry)

DE Four-copy gonadotropin-releasing hormone (GnRH) multimer.

GnRH multimer; gonadotropin-releasing hormone; immunosterilisation;
KW
KW immunosuppression; vaccine; feline; canine; equine; cervine; ds

na	
OS	Mammalia.
OS	Synthetic.

PN W09962545-A2.

PD 09-DEC-1999.

AA
PF 28-MAY-1999; 99WO-CA00493.

04 -JUN-1998: 98US-0088024.

PR 06-MAY-1999; 99US-0306689.

PA (BIOS-) BIOSTAR INC.

PI Robbins SC;

DR WPI: 2000-086857/07.

DR N-PSDB; AAZ55702.

Hormone immunogens, analogues or antibodies used to manufacture
PT vaccines for suppression of reproductive behavior and fertility in
PT vertebrates -

PS Claim 5; Fig 5B; 88pp; English.

This sequence represents a four-copy gonadotropin-releasing hormone (GnRH) multimer, where the second and fourth GnRH sequence have a His to Asp substitution at position 2 of the GnRH sequence. The invention relates to GnRH immunogens, analogues or antibodies that cross-react with endogenous GnRH of a vertebrate. A specifically claimed immunogenic fusion protein (AAV58361) comprises, in the N to C-terminal direction, a synthetic peptide sequence (AAV58364), an eight copy GnRH multimer (composed of two copies of the 4xGnRH multimer sequence of AAV58363), the LRT protein (which functions as a carrier protein), and a second eight copy GnRH multimer. The fusion protein may be used in a vaccine composition for prepubertal administration to a vertebrate subject to result in prolonged suppression of reproductive behaviour and/or fertility. GnRH immunogens, analogues or antibodies are used to manufacture a composition or vaccine for immunosterilisation or immunoc contraception of feline, canine, equine or cervine subjects. The vaccines are used to suppress reproductive behaviour and/or fertility for at least 10 months. The prepubertal administration results in a prolonged, long-term suppression of testicular development and/or function in males, or a prolonged, long-term suppression of ovarian development and/or function in females. The methods provide a viable and desirable alternative to surgical forms of sterilisation that are currently used.

AA	Sequence	49 AA;
SQ		

Query Match	36.8%	Score	97.5	DB	21	Length	49
Best Local Similarity	42.9%	Pred. No.	1.7e-05				
Matches	21	Conservative	2	Mismatches	11	Indels	15
		Gaps	1				

QY 2 HWSYGLRPGSSGSPSLDEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
 | | | | | | | | | | : | | | | | | |
pB 2 HWSYGLRPGS-----GSODWSYGLRPGSGSSHWSYGLRP 35

RESULT 15

userul as a vaccine for fertility control of domestic or farm animals.

```

YY      2 HWSYGLRPGSSGSPSLDEKKIAKMEKASSVFNVVNSSSGPSLHWSYGLRP 50
          |||||
          |::|

```



```

Best Local Similarity 56.1%; Pred. No. 0.00012;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNSSG 39
Db 284 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNSSG 324

RESULT 7
ID AAR13179
XX AAR13179 standard; Protein; 335 AA.
AC AAR13179;
DT 29-AUG-1991 (first entry)
DE NS1_81(NVDP)4RLfauth.
KW Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
KW hybrid; influenza virus; non-structural protein 1; fusion.
OS Plasmodium falciparum.
OS Influenza virus (A/PR/8/34).
XX FH Key Location/Qualifiers
XX FT Region 1..81
XX FT /label= N-terminal of NS1
XX FT /note= "Influenza virus nonstructural protein 1"
XX FT - Region 82..97
XX FT /label= immunodominant repeat region
XX FT Peptide 98..103
XX FT /note= "four variant tetrapeptide repeat units"
XX FT Region 104
XX FT /label= synthetic linker
XX FT /label= artifact
XX FT /note= "see comments"
XX FT Region 105..209
XX FT /label= AAs 19-123 of CS protein
XX FT /note= "Region I contg. flanking region less
XX FT signal sequence"
XX FT Region 210
XX FT /label= artifact
XX FT /note= "see comments"
XX FT Region 211..335
XX FT /label= AAs 288-412 of CS protein
XX FT /note= "Region II flanking region"
XX PN EP432965-A.
XX PD 19-JUN-1991.
XX PF 06-DEC-1990; 90EP-0313257.
XX PR 08-DEC-1989; 89US-0447746.
XX PA (SMIK ) SMITHKLINE BEECHAM.
XX PA (USSA ) US SEC OF THE ARMY.
XX PA (BIOM-) BIOMEDICAL RES INST.
XX PI Gross MS, Gordon DM, Hollingdale MR;
XX WPI; 1991-179771/25.
XX DR Polypeptide comprising immunogenic determinants from P falciparum
XX PT - for vaccine against malaria infection in humans.
XX PS Example 5; Page 11; 18pp; English.
XX CC The polypeptide is prepd. by genetic engineering of genes encoding
XX CC the P. falciparum circumsporozoite (CS) protein [Dane et al.,
XX CC Science 225 : 593 (1984)], and the influenza virus non-structural
XX CC protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
XX CC (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1

```

```

CC (NS1_81) is linked to a synthetic sequence encoding four repeat
CC units (the variant form) from the immunodominant region, which in
CC turn is linked via a synthetic sequence to DNA encoding Region I
CC contg. flanking region less the 18 AA signal region. This is
CC linked to DNA encoding Region II-contg. flanking region. The Pro
CC residue separating the Asp (at the C-terminal of the linker) from
CC the Region I-contg. CS flanking region is an artifact of a filled-
CC in BamHI site; the Gly separating the Region I and II-contg. CS
CC flanking regions is an artifact of a synthetic foki/TthIII I
CC linker. The peptide can be used in a vaccine for protection
CC against malaria.
CC See also AAR12306-R12311 and AAR13175-R13178.
XX SQ Sequence 335 AA;
XX
XX Query Match 37.4%; Score 99; DB 12; Length 335;
XX Best Local Similarity 56.1%; Pred. No. 0.00012;
XX Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;
XX
QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNSSG 39
Db 284 IKPGSANKPKDELVDYENDIEKKICKMEKCSSVFNVNSSG 324

RESULT 8
ID AAR83144
XX AAR83144 standard; protein; 411 AA.
AC AAR83144;
DT 20-NOV-1990 (first entry)
DE Sequence encoded by the circumsporozoite (CS) gene from
DE Plasmodium falciparum.
XX DE
XX DE Vaccine; antigen; immunogen; probe; hybridisation;
XX KW immunocassay; diagnosis.
XX KW
XX OS Plasmodium falciparum.
XX FH Key Location/Qualifiers
XX FT Region 106..120
XX FT /note= "Region I"
XX FT Region 147..206
XX FT /note= "Repeat region, repeat unit = NANP"
XX FT Region 123..146
XX FT /note= "Repeat region, repeat unit = NANPNVDP"
XX FT Region 211..286
XX FT /note= "Repeat region, repeat unit = NANP"
XX PN EP278940-A.
XX PD 17-AUG-1988.
XX PF 25-JAN-1988; 88EP-0870008.
XX PR 30-JAN-1987; 87US-0009325.
XX PA (SMIK ) SMITH KLINE-RIT.
XX PI Cabezon T, De Wilde M, Harford N;
XX WPI; 1988-229751/33.
XX DR N-PSDB; AAN81108.
XX PT DNA encoding hepatitis B virus antigens and hybrids contg. them -
XX PT used for expression in yeast to obtain vaccines and bivalent
XX PT vaccines
XX PS Example; Fig 3Aa-3Af; 101pp; English.
XX CC Sequence of the CS gene (AAN81108) is from lambda-mpfl. A recombinant
XX CC DNA molecule is claimed, comprising functional DNA coding sequence fused,

```

```

FT Region 88 /label= artifact
FT /note= "see comments"
FT Region 89..193 /label= AAs 19-123 of CS protein
FT /note= "Region I contg. flanking region less
FT signal sequence"
FT Region 194..201 /label= immunodominant repeat region
FT /note= "two tetrapeptide repeat units"
FT Region 202 /label= artifact
FT /note= "see comments"
FT Region 203..327 /label= AAs 288-412 of CS protein
FT /note= "Region II flanking region"
XX EP432965-A.
XX PN
XX XX
XX PD 19-JUN-1991.
XX PF 06-DEC-1990; 90EP-0313257.
XX PR 08-DEC-1989; 89US-0447746.
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF THE ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX Gross MS, Gordon DM, Hollingdale MR;
XX WPI; 1991-179771/25.
XX Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX Example 3; Page 10; 18pp; English.
XX The polypeptide is prepd. by genetic engineering of genes encoding
XX the P. falciparum circumsporozoite (CS) protein [Dane et al.,
XX Science 225 : 593 (1984)], and the influenza virus non-structural
XX protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
XX (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX (NS1-81) is linked via a synthetic sequence to DNA encoding Region
XX I contg. flanking region less the 18 AA signal region. This is
XX linked to a synthetic sequence encoding two repeat units from the
XX immunodominant region, which in turn is fused to DNA encoding
XX Region II-contg. flanking region. The pro residue separating the
XX Asp (at the C-terminal of the linker) from the Region I-contg. CS
XX flanking region is an artifact of a filled-in BamHI site; the Gly
XX separating the repeat units and the Region II-contg. CS flanking
XX region is an artifact of a synthetic FokI/TthIII I linker. The
XX peptide can be used in a vaccine for protection against malaria.
XX See also AAR12306-R12311 and AAR13175-R13179.
XX SQ Sequence 327 AA;

Query Match 37.4%; Score 99; DB 12; Length 327;
Best Local Similarity 56.1%; Pred. No. 0.00012;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPCSSGPSLD-----EKKIAKMEKASSVFNVNVSNG 39
:||||| | |||| |||| |||| |||| |||| ||||
Db 276 IKPGSANKPDELDYNDIEKKICKMEKCSVFNVSNG 316

RESULT 6
AAR13178
ID AAR13178 standard; Protein; 335 AA.
XX AAR13178;
XX AAR13178;
XX DT 29-AUG-1991 (first entry)

```

```

XX NS1_81(NANP)4RlfAuth.
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX hybrid; influenza virus; non-structural protein 1; fusion.
XX Plasmodium falciparum.
XX Influenza virus (A/PR/8/34).
XX Key Location/Qualifiers
XX Region 1..81 /label= N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX Region 82..97 /label= immunodominant repeat region
XX /note= "four tetrapeptide repeat units"
XX Peptide 98..103 /label= synthetic linker
XX Region 104 /label= artifact
XX /note= "see comments"
XX Region 105..209 /label= AAs 19-123 of CS protein
XX /note= "Region I contg. flanking region less
XX signal sequence"
XX Region 210 /label= artifact
XX /note= "see comments"
XX Region 211..335 /label= AAs 288-412 of CS protein
XX /note= "Region II flanking region"
XX EP432965-A.
XX PN
XX XX
XX PD 19-JUN-1991.
XX PF 06-DEC-1990; 90EP-0313257.
XX PR 08-DEC-1989; 89US-0447746.
XX (SMIK ) SMITHKLINE BEECHAM.
XX (USSA ) US SEC OF THE ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX Gross MS, Gordon DM, Hollingdale MR;
XX WPI; 1991-179771/25.
XX Polypeptide comprising immunogenic determinants from P falciparum
XX - for vaccine against malaria infection in humans.
XX Example 4; Page 11; 18pp; English.
XX The polypeptide is prepd. by genetic engineering of genes encoding
XX the P. falciparum circumsporozoite (CS) protein [Dane et al.,
XX Science 225 : 593 (1984)], and the influenza virus non-structural
XX protein 1 (NS1), [Baez et al., Nucleic Acids Research, 8 : 5845
XX (1980)]. The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX (NS1-81) is linked to a synthetic sequence encoding four repeat
XX units from the immunodominant region, which in turn is linked via
XX a synthetic sequence to DNA encoding Region I contg. flanking
XX region less the 18 AA signal region. This is linked to DNA
XX encoding Region II-contg. flanking region. The pro residue sep-
XX arating the Asp (at the C-terminal of the linker) from the Region
XX I-contg. CS flanking region is an artifact of a filled-in BamHI
XX site; the Gly separating the Region I and II-contg. CS flanking
XX regions is an artifact of a synthetic FokI/TthIII I linker. The
XX peptide can be used in a vaccine for protection against malaria.
XX See also AAR12306-R12311 and AAR13175-R13179.
XX SQ Sequence 335 AA;

Query Match 37.4%; Score 99; DB 12; Length 335;

```

```

XX PN EP398540-A.
XX PD 22-NOV-1990.
XX PF 01-MAY-1990; 90EP-0304720.
XX PR 03-MAY-1989; 89US-0346863.
XX PA (SMIK ) SMITHKLINE BEECHAM.
XX PI Gross MS, Young JF;
XX WPI; 1990-350299/47.
XX DR N-PSDB; AAQ06580.
XX PT New polypeptide used in malaria vaccine - comprises immunogenic
XX PT determinants from 1st and 2nd flanking regions of plasmodium
XX PT surface protein and intermediate repeat domain
XX PS Example 2; Page 11-12; 24pp; English.
XX CC The product is useful in preparation of vaccines for treatment and
XX CC prophylaxis of plasmodium sporozite infection. It may be easily
XX CC produced in large pure quantities from a transformed E.coli
XX CC expression system.
XX SQ Sequence 319 AA;

Query Match 37.4%; Score 99; DB 11; Length 319;
Best Local Similarity 56.1%; Pred. No. 0.00011;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKIKAKMERASSVFNVNSSG 39
Db 268 IKPGSANKPKDELVDYNDIEKKICKMEKCSSVFNVNSSIG 308

RESULT 4
AAR13176
ID AAR13176 standard; Protein; 319 AA.
AC AAR13176;
XX
XX 29-AUG-1991 (first entry)
DE NSL_81-RLFAuth.
XX
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX KW hybrid; influenza virus; non-structural protein 1; fusion.
XX OS Plasmodium falciparum.
XX OS Influenza virus (A/PR/8/34).
XX FH Key Location/Qualifiers
XX FT Region 1..81
XX FT Peptide /label= N-terminal of NS1
XX FT Region 82..87 /note= "see comments"
XX FT Region 88 /label= "Region 1 contg. flanking region less
XX FT Region 194 signal sequence"
XX FT Region /label= artifact
XX FT Region 89..193 /note= "see comments"
XX FT Region /label= AAs 19-123 of CS protein
XX FT Region /note= "Region I contg. flanking region less
XX FT Region 195..319 /label= AAs 288-412 of CS protein
XX FT Region /note= "Region II flanking region"

```

```

XX PN EP432965-A.
XX PD 19-JUN-1991.
XX PF 06-DEC-1990; 90EP-0313257.
XX PR 08-DEC-1989; 89US-0447746.
XX PA (SMIK ) SMITHKLINE BEECHAM.
XX PA (USSA ) US SEC OF THE ARMY.
XX PA (BIOM-) BIOMEDICAL RES INST.
XX PI Gross MS, Gordon DM, Hollingdale MR;
XX WPI; 1991-179771/25.
XX PT Polypeptide comprising immunogenic determinants from P falciparum
XX PT - for vaccine against malaria infection in humans.
XX PS Example 2; Page 10; 18pp; English.
XX CC The polypeptide is prepd. by genetic engineering of genes encoding
XX CC the P. falciparum circumsporozoite (CS) protein (Dame et al.,
XX CC Science 225 : 593 (1984)), and the influenza virus non-structural
XX CC protein 1 (NS1). (Baez et al., Nucleic Acids Research, 8 : 5845
XX CC (1980)). The DNA encoding the 1st 81 AAs of the N-terminal of NS1
XX CC (NS1-81) is linked via a synthetic sequence to DNA encoding Region
XX CC I contg. flanking regionless the 18 AA signal region, which in
XX CC turn is fused to DNA encoding Region II-contg. flanking region.
XX CC This CS fusion is designated RLFAuth. The Pro residue separating
XX CC the Asp (at the C-terminal of the linker) from RLFAuth is an arti-
XX CC fact of a filled in BamHI site; the Gly separating Region I and
XX CC Region II-contg. CS flanking regions is an artifact of a synthetic
XX CC FokI/TthIII I linker. The peptide can be used in a vaccine for
XX CC protection against malaria.
XX CC The complete nucleotide and AA sequences are given in EP-304720,
XX CC filed May 1, 1990.
XX CC See also AAR12306-R12311 and AAR13175-R13179.
XX SQ Sequence 319 AA;

Query Match 37.4%; Score 99; DB 12; Length 319;
Best Local Similarity 56.1%; Pred. No. 0.00011;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKIKAKMERASSVFNVNSSG 39
Db 268 IKPGSANKPKDELVDYNDIEKKICKMEKCSSVFNVNSSIG 308

RESULT 5
AAR13177
ID AAR13177 standard; Protein; 327 AA.
XX
XX AC AAR13177;
XX 29-AUG-1991 (first entry)
XX DE NSL_81-RLFAuth + (NANP)2.
XX
XX KW Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;
XX KW hybrid; influenza virus; non-structural protein 1; fusion.
XX OS Plasmodium falciparum.
XX OS Influenza virus (A/PR/8/34).
XX FH Key Location/Qualifiers
XX FT Region 1..81 /label= N-terminal of NS1
XX FT Peptide /note= "Influenza virus nonstructural protein 1"
XX FT Peptide 82..87 /label= synthetic linker

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 26.8077 Seconds
(without alignments)
211.311 Million cell updates/sec

Title: US-09-848-834A-20

Perfect score: 265

Sequence: 1 XHWSYGLRPGSSGSLDEKK.....NVNSSSGPSLHWSYGLRPX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
A_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127.5	48.1	33	15	AA62715
2	99	37.4	309	12	AA62715
3	99	37.4	319	11	AA62715
4	99	37.4	319	12	AA62715
5	99	37.4	327	12	AA62715
6	99	37.4	327	12	AA62715
7	99	37.4	335	12	AA62715
8	99	37.4	335	12	AA62715
9	99	37.4	411	9	AA62715
10	99	37.4	412	7	AA62715
11	97.5	36.8	49	17	AA62715

12	97.5	36.8	49	19	AAW79567	GnRH-2. Synthetic
13	97.5	36.8	49	19	AAW61542	Peptide hormone Gn
14	97.5	36.8	49	21	AAW58135	Four-copy gonadotr
15	97.5	36.8	49	21	AAW58135	GnRH analogue mult
16	97.5	36.8	544	17	AAW03943	LKT-GnRH protein f
17	97.5	36.8	544	19	AAW79570	LKT-GnRH chimeric
18	97.5	36.8	695	19	AAW79573	LKT-GnRH chimeric
19	97.5	36.8	695	21	AAW58133	Leukotoxin/gonadot
20	97.5	36.8	695	21	AAW58133	Gonadotropin relea
21	97.5	36.8	977	17	AAW03942	LKT-GnRH protein f
22	97.5	36.8	977	19	AAW79569	LKT-GnRH chimeric
23	95	35.8	350	21	AAW70278	Recombinant vaccin
24	95	35.8	412	9	AAW0835	Sequence encoded b
25	94	35.5	143	21	AAW49252	N6 polypeptide car
26	94	35.5	218	21	AAW49253	N10 polypeptide ca
27	94	35.5	240	21	AAW49254	N11 polypeptide ca
28	94	35.5	390	21	AAW49255	N19 polypeptide ca
29	91	34.3	424	14	AAW37796	RTS protein. Synt
30	90.5	34.2	42	21	AAW20865	GnRH tandem dimer
31	90	34.0	23	21	AAW20864	Peptide comprising
32	89.5	33.8	33	22	AAW63663	A peptide which ma
33	89.5	33.8	33	22	AAW63516	Vaccine related WH
34	89	33.6	19	22	AAW98951	Sequence of modifi
35	89	33.6	21	10	AAW91504	Malaria circumspor
36	89	33.6	21	16	AAW82586	P. falciparum CS p
37	89	33.6	21	16	AAW78920	Malaria circumspor
38	89	33.6	21	16	AAW75955	Circumsporozoite h
39	89	33.6	21	16	AAW70912	T-cell stimulatory
40	89	33.6	21	18	AAW05612	Peptide derived fr
41	89	33.6	21	20	AAW35440	Pathogen derived T
42	89	33.6	21	21	AAW80071	T helper cell (Th)
43	89	33.6	21	21	AAW54553	Unidentified pepti
44	89	33.6	21	21	AAW58777	
45	89	33.6	21	21	AAW58777	

ALIGNMENTS

RESULT 1

AA62715
ID AAR62715 standard; peptide; 33 AA.

XX AAR62715;

XX AAR62715;

DT 10-SEP-1995 (first entry)

XX LHRH-containing immunogenic peptide.

DE Helper T cell epitope; universal immune stimulator; Invasin; haptens;

KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;

KW androgen-dependent carcinoma; antitumour; infertility;

KW Plasmodium falciparum circumsporozoite.

XX Synthetic.

OS Synthetic.

XX Key Location/Qualifiers

FT Domain 1..21

FT FT /note= "Plasmodium falciparum circumsporozoite"

FT FT helper T cell epitope"

FT Domain 24..33

FT FT /note= "LHRH haptens"

XX WO9425060-A.

PN 10-NOV-1994.

PD NSI-81-RLFAuth. P

XX NSI-81-RLFAuth. + (

XX NSI-81(NANP)4RLFAu

XX NSI 81(NVDP)4RLFAu

XX Sequence encoded b

XX CS protein of mala

PR 27-APR-1993; 93US-0057166.

PR 14-APR-1994; 94US-0229275.

XX (LADD/) LADD A E.


```
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-844;
RA la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-844;
RA Jongwutiwes S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M83169; AAA29547.1; -.
DR EMBL; M83149; AAA29562.1; -.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
SQ SEQUENCE 424 AA; 45592 MW; F20CEB60636DB98E CRC64;

Query Match 37.4%; Score 99; DB 5; Length 424;
Best Local Similarity 56.1%; Pred. No. 0.00012;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNSSG 39
Db 373 IKPGSANKPKDELDYENDIEKKICKMEKCSVFNVNSSIG 413
:||||: | ||||| ||||| ||||| ||||| |

RESULT 14
Q99256 PRELIMINARY; PRT; 424 AA.
ID Q99256
AC Q99256;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN PRECURSOR (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91270295; PubMed=2052038;
RA Lockyer M.J.;
RT "Clonal variation in the Plasmodium falciparum circumsporozoite
RT protein gene.";
RL Mol. Biochem. Parasitol. 45:179-181(1991).
DR EMBL; M57499; AAA63422.1; -.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
KW Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 >424 CIRCUMSPOROZOITE PROTEIN.
FT NON_TER 424 424
SQ SEQUENCE 424 AA; 45610 MW; BF6D0F906C664B8F CRC64;

Query Match 37.4%; Score 99; DB 5; Length 424;
Best Local Similarity 56.1%; Pred. No. 0.00012;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNSSG 39
Db 373 IKPGSANKPKDELDYENDIEKKICKMEKCSVFNVNSSIG 413
:||||: | ||||| ||||| ||||| ||||| |
```

```
RESULT 15
Q27246 PRELIMINARY; PRT; 432 AA.
ID Q27246
AC Q27246;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-946;
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-946;
RA la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-946;
RA Jongwutiwes S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M83155; AAA29568.1; -.
DR EMBL; M83170; AAA29548.1; -.
DR EMBL; M83152; AAA29565.1; -.
DR EMBL; M83158; AAA29571.1; -.
DR EMBL; M83166; AAA29544.1; -.
DR EMBL; M83168; AAA29546.1; -.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
SQ SEQUENCE 432 AA; 46414 MW; 8787E6005578873A CRC64;

Query Match 37.4%; Score 99; DB 5; Length 432;
Best Local Similarity 56.1%; Pred. No. 0.00012;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNSSG 39
Db 381 IKPGSANKPKDELDYENDIEKKICKMEKCSVFNVNSSIG 421
:||||: | ||||| ||||| ||||| ||||| |

Search completed: October 10, 2002, 16:10:09
Job time : 19.8333 secs
```



```
DR PRINTS: PR01303; CRCMSPRZOITE.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12925 MW; 697DF63EEBDA90 CRC64;

Query Match 37.4%; Score 99; DB 5; Length 115;
Best Local Similarity 56.1%; Pred. No. 2.5e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNVSIG 39
Db 64 IKGSANKPKDELVDYEDIEKKICKMEKCSSVFNVNVSIG 104

RESULT 6
Q25835 ID Q25835 PRELIMINARY; PRT; 115 AA.
AC Q25835;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=828;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83157; AAA29570.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12925 MW; 9B52CFEF12BD48F CRC64;

Query Match 37.4%; Score 99; DB 5; Length 115;
Best Local Similarity 56.1%; Pred. No. 2.5e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNVSIG 39
Db 64 IKGSANKPKDELVDYEDIEKKICKMEKCSSVFNVNVSIG 104

RESULT 7
Q25836 ID Q25836 PRELIMINARY; PRT; 115 AA.
AC Q25836;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=834B;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83159; AAA29572.1; -.

DR InterPro; IPR003067; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12926 MW; 697DF630E50B3A90 CRC64;

Query Match 37.4%; Score 99; DB 5; Length 115;
Best Local Similarity 56.1%; Pred. No. 2.5e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNVSIG 39
Db 64 IKGSANKPKDELVDYEDIEKKICKMEKCSSVFNVNVSIG 104

RESULT 8
Q25839 ID Q25839 PRELIMINARY; PRT; 115 AA.
AC Q25839;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835C;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83162; AAA29575.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12974 MW; 6BB538287260DA90 CRC64;

Query Match 37.4%; Score 99; DB 5; Length 115;
Best Local Similarity 56.1%; Pred. No. 2.5e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNVSIG 39
Db 64 IKGSANKPKDQDLYEDIEKKICKMEKCSSVFNVNVSIG 104

RESULT 9
Q25794 ID Q25794 PRELIMINARY; PRT; 117 AA.
AC Q25794;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Doolan D.L., Saul A., Good M.F.;
RT "Geographically restricted heterogeneity of the Plasmodium falciparum
circumsporozoite protein: relevance for vaccine development.";
RT Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
RL
```

```

Query Match      41.5%; Score 110; DB 5; Length 436;
Best Local Similarity 61.0%; Pred. No. 4.3e-06;
Matches 25; Conservative 3; Mismatches 5; Indels 8;

Qy 7 LRGSGSPSLD-----EKKIAKMEKASSVFNVNSSG 39
:||||:| | | | | | | | | | | | | | | | | | | |
Db 385 IKPGSAGKPKDELYENDIEKKICKMEKCSSVFNVNSSIG 425

RESULT 4
Q25838 PRELIMINARY; PRT; 420 AA.
AC Q25838;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN.
GN GN
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835B;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
   falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83161; AAA29574.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPL1.1
DR SMART; SM00209; TSPL1.1
SQ SEQUENCE 420 AA; 4515 MW; 3A85B92432C2893C CRC64;

Query Match      39.6%; Score 105; DB 5; Length 420;
Best Local Similarity 58.5%; Pred. No. 1.9e-05;
Matches 24; Conservative 3; Mismatches 6; Indels 8;

Qy 7 LRGSGSPSLD-----EKKIAKMEKASSVFNVNSSG 39
:||||:| | | | | | | | | | | | | | | | | | | |
Db 369 IKPGSAGKPKDELYENDIEKKICKMEKCSSVFNVNSSIG 409

RESULT 5
Q9TVM7 PRELIMINARY; PRT; 115 AA.
AC Q9TVM7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CIRCUMSPOROZOITE PROTEIN (FRAGMENT).
GN GN
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=815, 947, AND 808;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
   falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83153; AAA29566.1; -.
DR EMBL; M83171; AAA29549.1; -.
DR EMBL; M83151; AAA29564.1; -.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF00090; tsp.1; 1.

```

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	110	41.5	117	5	Q25797	Q25797 plasmodium
2	110	41.5	408	5	Q25729	Q25729 plasmodium
3	110	41.5	436	5	Q27325	Q27325 plasmodium
4	105	39.6	420	5	Q25838	Q25838 plasmodium
5	99	37.4	115	5	Q9T9W7	Q9T9W7 plasmodium
6	99	37.4	115	5	Q25835	Q25835 plasmodium
7	99	37.4	115	5	Q25836	Q25836 plasmodium
8	99	37.4	115	5	Q25839	Q25839 plasmodium
9	99	37.4	117	5	Q25794	Q25794 plasmodium
10	99	37.4	117	5	Q25795	Q25795 plasmodium
11	99	37.4	117	5	Q25796	Q25796 plasmodium
12	99	37.4	393	5	Q99255	Q99255 plasmodium
13	99	37.4	424	5	Q27425	Q27425 plasmodium
14	99	37.4	424	5	Q99256	Q99256 plasmodium
15	99	37.4	432	5	Q27246	Q27246 plasmodium
16	99	37.4	432	5	Q25827	Q25827 plasmodium

ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-16

Query Match 36.8% Score 97.5; DB 3; Length 699;
Best Local Similarity 42.9%; Pred. No. 0.00021;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;
Qy 2 HWSYGLRPGSGPSLDEKKTAKMEKASVFNVNNSGPGSLHWSYGLRP 50
|||||
Db 646 HWSYGLRPGS-----GSQDWSYGLRPGSGSQHWSYGLRP 679

Search completed: October 10, 2002, 16:14:07
Job time : 11.7885 secs

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-124-491-10

Query Match          36.8%; Score 97.5; DB 3; Length 544;
Best Local Similarity 42.9%; Pred. No. 0.00015;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY      2 HWSYGRLPGSSGSLDEKKIAKMEKASSVFNVNVSNGSPSLHWSYGLRP 50
        ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      495 HWSYGLRPGS-----GSQDWSYGLRPGSGSQHWSYGLRP 528

RESULT 14
US-08-694-865-16
; Sequence 16, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/694,865
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-694-865-16

Query Match          36.8%; Score 97.5; DB 2; Length 699;
Best Local Similarity 42.9%; Pred. No. 0.00021;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY      2 HWSYGLRPGSSGSLDEKKIAKMEKASSVFNVNVSNGSPSLHWSYGLRP 50
        ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db      646 HWSYGLRPGS-----GSQDWSYGLRPGSGSQHWSYGLRP 679

RESULT 15
US-09-124-491-16
; Sequence 16, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
```

```

> TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
> NUMBER OF SEQUENCES:  34
> CORRESPONDENCE ADDRESS:
> ADDRESSEE:  REED & ROBINS LLP
> STREET:  285 HAMILTON AVENUE, SUITE 200
> CITY:  PALO ALTO
> STATE:  CA
> COUNTRY:  USA
> ZIP:  94301
> COMPUTER READABLE FORM:
> MEDIUM TYPE:  Floppy disk
> COMPUTER:  IBM PC compatible
> OPERATING SYSTEM:  PC-DOS/MS-DOS
> SOFTWARE:  PatentIn Release #1.0, Version #1.30
> CURRENT APPLICATION DATA:
> APPLICATION NUMBER:  US/08/694,865
> FILING DATE:  09-AUG-1996
> CLASSIFICATION:  424
> ATTORNEY/AGENT INFORMATION:
> NAME:  MCCracken, THOMAS P.
> REGISTRATION NUMBER:  38,548
> REFERENCE/DOCKET NUMBER:  9001-0016.22
> TELECOMMUNICATION INFORMATION:
> TELEPHONE:  (415)327-3400
> TELEFAX:  (415)327-3231
> INFORMATION FOR SEQ ID NO: 10:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 544 amino acids
> TYPE: amino acid
> TOPOLOGY: linear
> MOLECULE TYPE: protein
> US-08-694-865-10

Query Match          36.8%; Score 97.5; DB 2; Length 544;
Best Local Similarity 42.9%; Pred. NO. 0.00015;
Matches 21; Conservative 2; Mismatches 11; Indels 1

Qy  2 HWSYGLRPGSGPSLDEKKIAKMEKASVFNVNNSSGPSLHWSYGLRP 50
      |||||
Db  495 HWSYGLRPGS-----GSDHWSYGLRPGSGSHWSYGLRP 528

RESULT 12
US-08-878-748-10
> Sequence 10, Application US/08878748
> Patent No. 5969126
> GENERAL INFORMATION:
> APPLICANT:  POTTER, ANDREW A.
> APPLICANT:  REDMOND, MARK J.
> APPLICANT:  HUGHES, HOW P.A.
> TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
> NUMBER OF SEQUENCES:  28
> CORRESPONDENCE ADDRESS:
> ADDRESSEE:  REED & ROBINS
> STREET:  635 BRYANT STREET
> CITY:  PALO ALTO
> STATE:  CALIFORNIA
> COUNTRY:  UNITED STATES OF AMERICA
> ZIP:  94301
> COMPUTER READABLE FORM:
> MEDIUM TYPE:  Floppy disk
> COMPUTER:  IBM PC compatible
> OPERATING SYSTEM:  PC-DOS/MS-DOS
> SOFTWARE:  PatentIn Release #1.0, Version #1.25
> CURRENT APPLICATION DATA:
> APPLICATION NUMBER:  US/08/878,748
> FILING DATE:  19-JUN-1997
> CLASSIFICATION:  536
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER:  US 08/387,156
> FILING DATE:  10-FEB-1995
> APPLICATION NUMBER:  US 07/960,932
> FILING DATE:  14-OCT-1992

```


STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-4

Query Match 36.8%; Score 97.5; DB 2; Length 49;
Best Local Similarity 42.9%; Pred. No. 7.7e-06;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

Qy 2 HWSYGLRPGSSGSLDEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
|||||
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 35

RESULT 8
US-08-878-748-4
Sequence 4, Application US/08878748
Patent No. 5969126
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HOW P.A.
TITLE OF INVENTION: GPRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US/07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:

NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-4

Query Match 36.8%; Score 97.5; DB 2; Length 49;
Best Local Similarity 42.9%; Pred. No. 7.7e-06;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

Qy 2 HWSYGLRPGSSGSLDEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
|||||
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 35

RESULT 9
US-09-124-491-4
Sequence 4, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GPRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-4

REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-797A-3

Query Match 37.4%; Score 99; DB 2; Length 424;
Best Local Similarity 56.1%; Pred. No. 7.1e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

Qy 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNSSSG 39
Db 152 IKPGSANKPKDELVDYNDIEKKICKMERKCSSVFNVNSSIG 192

RESULT 5

US-08-932-929B-3
; Sequence 3, Application US/08932929B
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HbSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932.929B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,797
; FILING DATE: 04-DEC-1996
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1FWC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-932-929B-3

Query Match 37.4%; Score 99; DB 4; Length 424;
Best Local Similarity 56.1%; Pred. No. 7.1e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

Qy 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNSSSG 39
Db 152 IKPGSANKPKDELVDYNDIEKKICKMERKCSSVFNVNSSIG 192

RESULT 6

US-08-387-156-4
; Sequence 4, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-4

Query Match 36.8%; Score 97.5; DB 1; Length 49;
Best Local Similarity 42.9%; Pred. No. 7.7e-06;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

Qy 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
Db 2 HWSYGLRPGS-----GSDWSYGLRPGSSQHSYGLRP 35

RESULT 7

US-08-694-865-4
; Sequence 4, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: REED & ROBINS LLP

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L32864; AAA31066.1; -;
 DR PIR; A01411; RHPGG;
 DR InterPro; IPR002012; GNRH;
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBRN1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 91 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

Query Match 29.0%; Score 54; DB 1; Length 91;
 Best Local Similarity 90.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 24 SLHWSYGLRP 33
 | | | | | | | |
 Db 23 SQHWSYGLRP 32

Search completed: October 10, 2002, 16:06:29
 Job time : 5.43162 secs

```

DR PIR: A35396; A35396.
DR PIR: A43503; A43503.
DR HSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_Mtpeptidse.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 1 448 BOTULINUM NEUROTOXIN C1, LIGHT-CHAIN.
FT CHAIN 449 1290 BOTULINUM NEUROTOXIN C1, HEAVY-CHAIN.
FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 436 452 INTERCHAIN (PROBABLE).
FT CONFLICT 84 84 P -> T (IN REF. 2).
SQ SEQUENCE 1290 AA; 148734 MW; 71FBE379F97129E8 CRC64;

Query Match 29.3%; Score 54.5; DB 1; Length 1290;
Best Local Similarity 30.8%; Pred. NO. 13;
Matches 12; Conservative 10; Mismatches 8; Indels 9; Gaps 2;

QY 1 FNNFTVSFWLRVPKYSASHLEGPSL-----HMSYGL 31
DB 934 YESFSISFWIRINK-WVSNLPGYTIIDSVKNNSGWSIGI 971

RESULT 14
GONI_MACMU
ID GONI_MACMU STANDARD; PRT; 67 AA.
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor (Contains: Gonadoliberein I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques."
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S75918; AAB33096.1; -.
DR InterPro: IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.

```

```

DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Signal.
FT NON_TER 1 1
FT SIGNAL <1 5 BY SIMILARITY.
FT CHAIN 6 >67 PRONADOLIBERIN I.
FT PEPTIDE 6 15 GONADOLIBERIN I.
FT PEPTIDE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
SIMILARITY).
FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
SIMILARITY).
FT NON_TER 67 67
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 29.0%; Score 54; DB 1; Length 67;
Best Local Similarity 90.0%; Pred. NO. 0.82;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 SLHWSYGLRP 33
DB 5 SQHWSYGLRP 14

RESULT 15
GONI_PIG
ID GONI_PIG STANDARD; PRT; 91 AA.
AC P49921;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor (Contains: Gonadoliberein I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].
GN GNRH1 OR GNRH.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Hypothalamus;
RA Wessner G.D., Matteri R.L., Becker B.A.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 24-33.
RX MEDLINE=72114303; PubMed=4946067;
RA Baba Y., Matsuo H., Schally A.V.;
RT "Structure of the porcine LH- and FSH-releasing hormone. II.
RT Confirmation of the proposed structure by conventional sequential
RT analyses."
RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN [3]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE=72063376; PubMed=4942726;
RA Matsuo H., Azimura A., Nair R.M.G., Schally A.V.;
RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RT phase method."
RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN [4]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE=72117544; PubMed=4946275;
RA Baba Y., Arimura A., Schally A.V.;
RT "On the tryptophan residue in porcine LH and FSH-releasing hormone."
RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----

```

RN [3] SEQUENCE FROM N.A.
RX MEDLINE=85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
releasing hormone.";
RL Nature 311:666-668(1984).
RN [4]
RX SEQUENCE OF 24-33.
RA MEDLINE=83126573; PubMed=6760865;
RT Tan L., Rousseau P.;
"The chemical identity of the immunoreactive LHRH-like peptide
biosynthesized in the human placenta.";
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
HORMONES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
Lutrepulse or Lutrel (Ferring Pharmaceuticals) and Relisorm
(Serono).
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01059; CAA29526.1; -
DR EMBL; M12578; AAA35916.1; -
DR EMBL; X15215; CAA33285.1; -
DR PIR; A01410; RHUG.
DR PIR; A26173; A26173.
DR PIR; S05308; S05308.
DR MIM; 152760; -
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; Gonadolibertin.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBRN.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Placenta; Pharmaceutical; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT CONFLICT 16 16 W -> S (IN REF. 3).
SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;
Query Match 29.3%; Score 54.5; DB 1; Length 92;
Best Local Similarity 73.3%; Pred. No. 0.96;
Matches 11; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 20 LEG-PSLHWSYGLRP 33
Db 18 LEGCSQHSYGLRP 32
: || | ||||| |||
RESULT 13
BXCL_CLOBO STANDARD; PRT; 1290 AA.
ID BXCL_CLOBO
AC P18640;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BONT/C1)

DE (Bontolysin Cl).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=90370487; PubMed=2204031;
RA Hauser D., Eklund M.W., Kurazona H., Binz T., Niemann H., Gill D.M.,
Raquet P., Popoff M.R.;
RT "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RL Nucleic Acids Res. 18:4924-4924(1990).
RN [2]
RN SEQUENCE FROM N.A.
RX STRAIN=TYPE C STOCKHOLM / C-ST;
RX MEDLINE=91024998; PubMed=2222445;
RA Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,
Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT "The complete nucleotide sequence of the gene coding for botulinum
type C1 toxin in the C-ST phage genome.";
RL Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
RN [3]
RN SEQUENCE OF 2-25.
RX STRAIN=TYPE C STOCKHOLM / C-ST;
RX MEDLINE=8813072; PubMed=2450068;
RA Tsuzuki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,
Oguma K.;
RT "Establishment of a monoclonal antibody recognizing an antigenic site
common to Clostridium botulinum type B, C1, D, and E toxins and
tetanus toxin.";
RL Infect. Immun. 56:898-902(1988).
RN [4]
RN IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94038966; PubMed=7901002;
RA Blasi J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;
RT "Botulinum neurotoxin C1 blocks neurotransmitter release by means of
cleaving HPC-1/syntaxin.";
RL EMBO J. 12:4821-4828(1993).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
detected action on small molecule substrates.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
BOTULINUS NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -!- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C
STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
BACTERIOPHAGE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66433; CAA47060.1; -
DR EMBL; X72793; CAA51313.1; -
DR EMBL; X53751; CAA37780.1; -
DR EMBL; D90210; BAA14235.1; -
DR EMBL; X62389; CAA44263.1; -
DR PIR; S11291; S11291.

```

FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 212 212 BY SIMILARITY.
FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT FT METAL 411 425 INTERCHAIN (PROBABLE).
FT CONFLICT 229 229 K -> M (IN REF. 2).
SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C2312857 CRC64;

Query Match 30.4%; Score 56.5; DB 1; Length 1250;
Best Local Similarity 22.4%; Pred. No. 7;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

QY 1 FNNFTVSFWLRVP-----KVSASHLE-----GP 23
Db 911 YKNFSISFWPRIPYNDKNIVNVNNEYTIINCRDNNSGWKVSLNHNHETIWTLDNSGINQ 970
QY 24 SLHWSYG 30
Db 971 KLFNYG 977

RESULT 11
BXA2_CLOBO STANDARD; PRT; 1295 AA.
AC Q45894; P77780;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kyoto-F;
RX MEDLINE=94143603; PubMed=8310180;
RA Willens A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium
botulinum type A associated with infant botulism: comparison with
other clostridial neurotoxins.";
RL Res. Microbiol. 144:547-556(1993).
RN [2]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=Kyoto-F;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
components of the botulinum toxin complex in proteolytic Clostridium
botulinum types A, B, and F: evidence of chimeric sequences in the
gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
binds with high affinity to peripheral neuronal presynaptic
membrane, is then internalized by receptor-mediated endocytosis.
The C-terminus of the heavy chain (H) is responsible for the
adherence of the toxin to the cell surface while the N-terminus
mediates transport of the light chain from the endocytic vesicle
to the cytosol. After translocation, the light chain (L)
hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking
neurotransmitter release. Inhibition of acetylcholine release
results in flaccid paralysis, with frequent heart or respiratory
failure (By similarity).
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
detected action on small molecule substrates.
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
heavy chain (H) (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of

```

```

CC botulinum neurotoxin: Types A, B, Cl, D, E, F, and G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X73423; CAA51824.1; -
DR EMBL: X87974; CAA61234.1; -
DR HSP: P10845; 3BTA.
DR MEROPS; M27.002; -
DR InterPro; IPR000395; Bontoxilysin.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0 BY SIMILARITY.
FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 223 223 BY SIMILARITY.
FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 429 453 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1234 1279 BY SIMILARITY.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675 POTENTIAL.
SQ SEQUENCE 1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;

Query Match 30.1%; Score 56; DB 1; Length 1295;
Best Local Similarity 50.0%; Pred. No. 8.5;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
Db 937 YENFSTFWIKPK 950
: ||: ||: ||: ||: ||

RESULT 12
GONI_HUMAN STANDARD; PRT; 92 AA.
ID GONI_HUMAN
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Progonaoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
(luteinizing hormone releasing hormone I) (Gonadotropin releasing
hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
peptide I].
GN GNRH1 OR GNRH OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89366682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
hormone gene.";
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
of gonadotropin-releasing hormone and prolactin release-inhibiting
factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).

```

```

FT CHAIN 22 90 PROGNADOLIBERIN I.
FT PEPTIDE 22 31 GONADOLIBERIN I.
FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 31 31 IMCUTATION (G-32 PROVIDE AMIDE GROUP).
SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 30.9%; Score 57.5; DB 1; Length 90;
Best Local Similarity 80.0%; Pred. No. 0.37;
Matches 12; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Oy 20 LEG-PSLHWSYGLRP 33
   ||| | |||||
Db 16 LEGGSSQHSYGLRP 30

RESULT 9
VP2_AHSV6 STANDARD; PRT; 1051 AA.
AC O71024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Outer capsid protein VP2.
GN S2 OR L2.
OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
   (serotype 6)).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
ON NCBI_TaxID=86060;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98278331; PubMed=9617769;
RA Williams C.F., Inoue T., Lucas A.-M., Zanotto P., Roy P.;
RT "The complete sequence of four major structural proteins of African
   horse sickness virus serotype 6: evolutionary relationships within
   and between the orbiviruses.";
RL Virus Res. 53:53-73(1998).
CC -!- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
   WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
   MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF021235; AAC40994.1; -.
DR InterPro; IPR001742; Orbi_VP2.
DR Pfam; PF00898; Orbi_VP2; 2.
KW Coat protein.
SQ SEQUENCE 1051 AA; 122326 MW; 2B04DB9E389F4B5F CRC64;

Query Match 30.6%; Score 57; DB 1; Length 1051;
Best Local Similarity 47.6%; Pred. No. 5;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Oy 1 FNNFTVSFWLRVPKVSASHLE 21
   | : ||| | : |||
Db 636 FSKRVSYSYWRVEKITTKHLE 656

RESULT 10
BXE_CLOBU
ID BXE_CLOBU STANDARD; PRT; 1250 AA.
AC P30995;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)

```

```

DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
DE (Bontoxilysin E).
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
ON NCBI_TaxID=1492;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 43181, AND ATCC 43755;
RX MEDLINE=92181428; PubMed=1543481;
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulinum neurotoxin E derived from Clostridium
   botulinum type E (strain Beluga) and Clostridium butyricum (strains
   ATCC 43181 and ATCC 43755).";
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN [2]
RP SEQUENCE OF 1-251 FROM N.A.
RC STRAIN=BL6340;
RC MEDLINE=91237316; PubMed=2033376;
RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA Yokosawa N., Yashiki T., Oguma K.;
RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
   type E toxin gene from Clostridium butyricum strain BL6340.";
RL J. Gen. Microbiol. 137:519-525(1991).
RN [3]
RP SEQUENCE OF 1-48.
RC STRAIN=5262;
RA Glimenez J., Foley J., Dasgupta B.R.;
RT "Neurotoxin type E from Clostridium botulinum and C. butyricum;
   partial sequence and comparison.";
RL FASEB J. 2:A1750-A1750(1988).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
   RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
   AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
   WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
   INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
   ENDOPEPTIDASE.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
   neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
   detected action on small molecule substrates.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
   HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
   WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
   FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
   BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62088; CAA43998.1; -.
DR EMBL; X53180; CAA37321.1; -.
DR PIR; JH0256; JH0256.
DR PIR; S16145; S16145.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTpeptidse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0
FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.

```

DR	ProDom; PD001963; Bontoxilysin; 1.
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.
KW	Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT	INIT_MET 0
FT	CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
FT	CHAIN 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
FT	METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE 230 230 BY SIMILARITY.
FT	METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT	DISULFID 436 445 INTERCHAIN (PROBABLE).
FT	CONFLICT 29 29 T -> M (IN REF. 4).
FT	CONFLICT 217 217 R -> G (IN REF. 2).
FT	CONFLICT 224 224 A -> S (IN REF. 4).
FT	CONFLICT 463 463 S -> R (IN REF. 4).
SQ	SEQUENCE 1290 AA; 150670 MW; D21745E2C024DF43 CRC64;
Query Match 31.2%; Score 58; DB 1; Length 1290;	
Best Local Similarity 64.3%; Pred. No. 4.5;	
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;	
QY	1 FNNFTVSFWLRVPK 14 : : : :
Db	922 FLDFSFSFWIRPK 935
RESULT 8	
GONL_MOUSE	STANDARD; PRT; 90 AA.
ID	GONL_MOUSE
AC	P13562; DT 01-JAN-1990 (Rel. 13, Created)
DT	01-JAN-1990 (Rel. 13, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)]
DE	(Luteinizing hormone releasing hormone I) (Gonadotropin releasing hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor I].
DE	GNRH1 OR GNRH.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=87069928; PubMed=3024317;
RA	Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA	Phillips H.S., Nikolics K., Seeburg P.H.;
RT	"A deletion truncating the gonadotropin-releasing hormone gene is responsible for hypogonadism in the hpg mouse.";
RL	Science 234:1366-1371(1986).
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.
CC	-1- SUBCELLULAR LOCATION: Secreted.
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; M14872; AAA3717.1; -.
DR	MGD; MG1:95789; GnRh.
DR	InterPro; IPR002012; GNRH.
DR	InterPro; IPR004079; GonadoliberinI.
DR	Pfam; PF00446; GnRh; 1.
DR	PRINTS; PR01541; GONADOLIBERNI.
DR	PROSITE; PS00473; GNRH; 1.
KW	Cleaveage on pair of basic residues; Hormone; Amidation; Hypothalamus; placenta; Signal.
FT	SIGNAL 1 21

Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
 "The complete amino acid sequence of the Clostridium botulinum type-E
 neurotoxin, derived by nucleotide-sequence analysis of the encoding
 gene."; Eur. J. Biochem. 204:657-667(1992).
 [3]
 RN SEQUENCE OF 1-251 FROM N.A.
 RP MEDLINE=90264400; PubMed=2160960;
 RX MEDLINE=90264400; PubMed=2160960;
 RA Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
 "The complete sequence of botulinum neurotoxin type A and comparison
 with other clostridial neurotoxins."; J. Biol. Chem. 265:9153-9158(1990).
 [4]
 RN SEQUENCE OF 1-13.
 RP MEDLINE=85197963; PubMed=3888113;
 RX MEDLINE=85197963; PubMed=3888113;
 RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;
 "Partial amino acid sequences of botulinum neurotoxins types B and
 E."; Arch. Biochem. Biophys. 238:544-548(1985).
 [5]
 RN SEQUENCE OF 419-426.
 RP MEDLINE=90344918; PubMed=2116911;
 RX MEDLINE=90344918; PubMed=2116911;
 RA Gimenez J.A., Dasgupta B.R.;
 "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
 reveals the site trypsin nicks and homology with tetanus
 neurotoxin."; Biochimie 72:213-217(1990).
 [6]
 RN IDENTIFICATION OF SUBSTRATE.
 RP MEDLINE=94063091; PubMed=8243676;
 RX MEDLINE=94124495; PubMed=82944407;
 RA Binz T., Blas J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 Jahn R., Niemann H.;
 "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins."; J. Biol. Chem. 269:1617-1620(1994).
 [7]
 RN FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
 181 BOND IN SNAP-25.
 CC CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC SUBCELLULAR LOCATION: Secreted.
 CC MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X62089; CAA43999.1; -;
 DR EMBL; X62683; CAA44558.1; -;
 DR PIR; A60027; A60027.
 DR PIR; B35294; B35294.

PIR; JH0257; JH0257.
 DR PIR; S08575; S08575.
 DR PIR; S18111; S18111.
 DR PIR; S21178; S21178.
 DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -.
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_MTPeptidse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT_MET 0
 FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT METAL 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
 FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 411 425 INTERCHAIN (PROBABLE).
 FT CONFLICT 176 197 R -> G (IN REF. 2).
 FT CONFLICT 197 197 C -> S (IN REF. 2 AND 3).
 FT CONFLICT 339 339 R -> A (IN REF. 2).
 FT CONFLICT 772 772 I -> L (IN REF. 2).
 FT CONFLICT 962 963 FE -> LQ (IN REF. 2).
 FT CONFLICT 966 966 R -> A (IN REF. 2).
 FT CONFLICT 1194 1194 N -> NN (IN REF. 2).
 FT CONFLICT 1194 1194 N -> NN (IN REF. 2).
 SQ SEQUENCE 1250 AA; 143712 MW; D9FCE26DDA041EB4 CRC64;
 Query Match 31.5%; Score 58.5; DB 1; Length 1250;
 Best Local Similarity 23.6%; Pred. No. 3.7; Gaps 2;
 Matches 13; Conservative 9; Mismatches 4; Indels 29;
 Oy 1 FNNFTVSWLRVP-----KVSASHLEGPSLHWSY 29
 : ||::||::||: |
 Db 911 YKNFSISFWVRIPNVDNKIVNVNNEYTIINCNRDNGSKVSLNHNE---IIWTF 962
 : ||::||::||: |
 RESULT 7
 BXB_CLOBO STANDARD; PRT; 1290 AA.
 ID BXB_CLOBO
 AC P10844; P10843;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1993 (Rel. 25, Last sequence update)
 DE 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BONT/B)
 DE (Bontoxilysin B).
 GN BOTB.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RX MEDLINE=92384550; PubMed=1514783;
 RP SEQUENCE FROM N.A.
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
 RA Minton N.P.;
 RA "Molecular cloning of the Clostridium botulinum structural gene
 RA encoding the type B neurotoxin and determination of its entire
 RA nucleotide sequence."; Appl. Environ. Microbiol. 58:2345-2354(1992).
 RL [2]
 RN SEQUENCE OF 35-245 FROM N.A.
 RP STRAIN=NCTC 7273;
 RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 [3]
 RN SEQUENCE OF 633-993 FROM N.A.
 RP STRAIN=NCTC 7273;
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RA "Gene probes for identification of the botulinum neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";

Arch. Biochem. Biophys. 266:142-151(1988).
 [10] SEQUENCE OF 448-482.
 RP MEDLINE=85285016; PubMed=3896784;
 RA Shone C.C., Hambleton P., Melling J.;
 RA "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
 RT and purification of two tryptic fragments. Proteolytic action near
 RT the COOH-terminus of the heavy subunit destroys toxin-binding
 RT activity.";
 RL Eur. J. Biochem. 151:75-82(1985).
 [11] IDENTIFICATION OF SUBSTRATE.
 RP MEDLINE=94063091; PubMed=8243676;
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RA "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103(1993).
 [12] IDENTIFICATION OF SUBSTRATE.
 RP MEDLINE=94124495; PubMed=8294407;
 RA Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RA "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 [13] MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
 RP PubMed=11700044;
 RA Rigoni M., Caccini P., Johnson E.A., Montecucco C., Rossetto O.;
 RA "Site-directed mutagenesis identifies active-site residues of the
 RT light chain of botulinum neurotoxin type A.";
 RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
 [14] X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
 RP MEDLINE=98455071; PubMed=9783750;
 RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
 RA "Crystal structure of botulinum neurotoxin type A and implications
 RT for toxicity.";
 RL Nat. Struct. Biol. 5:898-902(1998).
 CC -I- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
 CC binds with high affinity to peripheral neuronal presynaptic
 CC membrane, is then internalized by receptor-mediated endocytosis.
 CC The C-terminus of the heavy chain (H) is responsible for the
 CC adherence of the toxin to the cell surface while the N-terminus
 CC mediates transport of the light chain from the endocytic vesicle
 CC to the cytosol. After translocation, the light chain (L)
 CC hydrolyzes the 197-Gln-I-Arg-198 bond in SNAP-25, thereby blocking
 CC neurotransmitter release. Inhibition of acetylcholine release
 CC results in flaccid paralysis, with frequent heart or respiratory
 CC failure.
 CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
 CC detected action on small molecule substrates.
 CC -I- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H).
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- PHARMACEUTICAL: Available under the name BOTOX(R) (Allergan) for
 CC the treatment of strabismus and blepharospasm associated with
 CC dystonia and cervical dystonia. Also used for the treatment of
 CC hemifacial spasm and a number of other neurological disorders
 CC characterized by abnormal muscle contraction.
 CC -I- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X52066; CAA36289.1; -;
 DR EMBL; M30196; AAA23262.1; -;
 DR EMBL; X92973; CAA63551.1; -;
 DR EMBL; D67030; BAA11051.1; -;
 DR EMBL; M27892; AAA23269.1; -;
 DR PIR; A35294; BTCLAB;
 DR PIR; S09492; S09492;
 DR PDB; 3BTA; 01-OCT-99;
 DR MEROPS; M27.002; -;
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_Mtpeptidse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR PRODOM; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
 KW Pharmacological; 3D-structure.
 FT INIT_MET 0 0
 FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
 FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
 FT METAL 222 222 ZINC (CATALYTIC).
 FT ACT_SITE 223 223 ZINC (CATALYTIC).
 FT METAL 226 226 ZINC (CATALYTIC).
 FT METAL 261 261 ZINC (CATALYTIC).
 FT DISULFID 429 453 INTERCHAIN.
 FT DISULFID 1234 1279
 FT TRANSMEM 626 646 POTENTIAL.
 FT TRANSMEM 655 675 POTENTIAL.
 FT VARIANT 26 26 V -> A.
 FT MUTAGEN 261 261 E->A: DRASTIC DECREASE IN ENZYMIC ACTIVITY.
 FT MUTAGEN 265 265 F->A: DECREASE IN ENZYMIC ACTIVITY.
 FT MUTAGEN 365 365 Y->A: DECREASE IN ENZYMIC ACTIVITY.
 FT CONFLICT 1 1 P -> Q (IN REF. 1).
 FT CONFLICT 479 479 E -> P (IN REF. 9).
 FT CONFLICT 875 875 T -> L (IN REF. 8).
 FT CONFLICT 891 891 S -> K (IN REF. 8).
 SQ SEQUENCE 1295 AA; 149322 MW; 858342F754862579 CRC64;
 Query Match 31.7%; Score 59; DB 1; Length 1295;
 Best Local Similarity 57.1%; Pred. No. 3.3;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 FNNETVSFWLRVVK 14
 : ||: |||: |||
 Db 937 YENFSTFWIRPK 950
 RESULT 6
 BXE_CLOBO
 ID BXE_CLOBO STANDARD; PRT; 1250 AA.
 AC Q00496;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
 DE (Bontoxilysin E).
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BEUGA;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT "Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92174922; PubMed=1541280;

```

RP SEQUENCE FROM N.A.
RC STRAIN=113 / 30;
RX MEDLINE=94092745; PubMed=8268233;
RA Campbell K., Collins M.D., East A.K.;
RT "Nucleotide sequence of the gene coding for Clostridium botulinum
  (Clostridium argentinense) type G neurotoxin: genealogical comparison
  with other clostridial neurotoxins.";
RL Biochim. Biophys. Acta 1216:487-491(1993).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
  RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
  AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
  WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
  INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
  ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
  neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
  detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
  HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
  WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
  FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
  BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X74162; CAA52275.1; -
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MPeptidse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PRODom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0 BY SIMILARITY.
FT CHAIN 1 441 BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT CHAIN 442 1296 BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 230 230 BY SIMILARITY.
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 435 449 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1296 AA; 149013 MW; DCBE47E15F665C31 CRC64;

Query Match 32.8%; Score 61; DB 1; Length 1296;
Best Local Similarity 38.1%; Pred. No. 1.8;
Matches 8; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

Oy 1 FNNFTVSWLRVPKVSASHLE 21
Db 929 FDNFSINFWRTPKYNNNDIQ 949

RESULT 5
BXAL_CLOBO STANDARD; PRT; 1295 AA.
AC P10845; P18639; P01561;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
  chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.

```

```

OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 2916;
RX MEDLINE=90233864; PubMed=2185020;
RA Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.-J.,
  Shone C.C., Atkinson T., Melling J., Minton N.P.;
RT "The complete amino acid sequence of the Clostridium botulinum type A
  neurotoxin, deduced by nucleotide sequence analysis of the encoding
  gene.";
RT Eur. J. Biochem. 189:73-81(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=62A;
RX MEDLINE=90264400; PubMed=2160960;
RA Biz B., Kuarzonzo H., Wille M., Frevent J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
  with other clostridial neurotoxins.";
RT J. Biol. Chem. 265:9153-9158(1990).
RN [3]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=62A;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
  components of the botulinum toxin complex in proteolytic Clostridium
  botulinum types A, B, and F: evidence of chimeric sequences in the
  gene encoding the nontoxic nonhemagglutinin component.";
RT Int. J. Syst. Bacteriol. 46:1105-1112(1996).
RN [4]
RP SEQUENCE OF 1-34 FROM N.A.
RC STRAIN=HALL;
RX MEDLINE=89350959; PubMed=2669749;
RA Betley M.J., Somers E., Dasgupta B.R.;
RT "Characterization of botulinum type A neurotoxin gene: delineation of
  the N-terminal encoding region.";
RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
RN [5]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=TYPE A NIH;
RX MEDLINE=96096783; PubMed=8521962;
RA Fujita R., Fujiwaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
  components of Clostridium botulinum type A progenitor toxins.";
RL FEBS Lett. 376:41-44(1995).
RN [6]
RP SEQUENCE OF 1-16.
RX MEDLINE=84178501; PubMed=6370252;
RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequence of the heavy and light chains of
  botulinum neurotoxin type A.";
RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN [7]
RP SEQUENCE OF 1-46.
RA Dasgupta B.R., Foley J., Niece R.;
RT "Partial sequence of the light chain of botulinum neurotoxin type A.";
RL Biochemistry 26:4162-4162(1987).
RN [8]
RP SEQUENCE OF 1-5 AND 444-456.
RX MEDLINE=91120847; PubMed=2126206;
RA Dasgupta B.R., Dekleva M.L.;
RT "Botulinum neurotoxin type A: sequence of amino acids at the
  N-terminus and around the nicking site.";
RL Biochimie 72:661-664(1990).
RN [9]
RP SEQUENCE OF 448-464 AND 872-895.
RX MEDLINE=89024662; PubMed=3178218;
RA Sartymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two
  halves and their partial sequences.";

```


EMBO J. 11:3577-3583(1992).
 [7]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 [8]
 RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RP MEDLINE=97475217; PubMed=9334741;
 RX Umland T.C., Wingert L.M., Swannathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792(1997).
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
 CC BOND OF SYNAPTOSOMAL-2.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-|-PHE-77 BOND IN
 CC SYNAPTOSOMAL-2.
 CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 CC AND ARE NON-TOXIC AFTER SEPARATION.
 CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 CC GANGLIOSIDE RECEPTORS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X04436; CAA28033.1; -;
 DR EMBL; M12739; AAA23282.1; -;
 DR EMBL; X06214; CAA29564.1; -;
 DR PIR; A25689; BTCLTN.
 DR PDB; 1AF9; 29-APR-98.
 DR PDB; 1A8D; 14-OCT-98.
 DR MEROPS; M27.001; -;
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_Mtpetdse.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR PRODOM; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure.
 FT INIT_MET 0 0
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 233 233 BY SIMILARITY.
 FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
 FT TRANSMEM 226 246 POTENTIAL.
 FT TRANSMEM 669 689 POTENTIAL.
 FT DISULFID 438 466 INTERCHAIN.
 FT DISULFID 1076 1092
 SQ SEQUENCE 1314 AA; 150550 MW; 134C3657133EF81D CRC64;
 Query Match 60.28; Score 112; DB 1; Length 1314;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FNNFTVFWLRVPRKVSASHLE 21

Db 946 FNNFTVFWLRVPRKVSASHLE 966
 RESULT 2
 GONI_RAT
 ID GONI_RAT STANDARD; PRT; 92 AA.
 AC P07490;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
 DE I].
 DE GNRH1 OR GNRH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89384661; PubMed=2476669;
 RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and
 RT hypothalamic expression.";
 RL Mol. Endocrinol. 3:1257-1262(1989).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93103480; PubMed=1468115;
 RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
 RT "Thymocytes express a mRNA that is identical to hypothalamic
 RT luteinizing hormone-releasing hormone mRNA.";
 RL Cell. Mol. Neurobiol. 12:447-454(1992).
 [4]
 RP SEQUENCE OF 1-47 FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=87149087; PubMed=3547652;
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
 RT "Two mammalian genes transcribed from opposite strands of the same
 RT DNA locus.";
 RL Science 235:1514-1517(1987).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES
 CC -1- SUBCELLULAR LOCATION: Secreted
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; S50870; AAB24572.1; -;
 DR EMBL; M12579; AAA41263.1; -;
 DR EMBL; M31670; AAA41264.1; -;
 DR EMBL; M15527; AAA42141.1; ALT_SEQ.
 DR EMBL; M15529; AAA42139.1; -;
 DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
 DR PIR; B26173; RHRTG.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 4.43162 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSEFWLRVPKVSASHLEGLHSWGLRXPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match	Length	ID	Description
1	112	60.2	1314	1	TETX_CLOTE	P04958 clostridium
2	61.5	33.1	92	1	GONL_RAT	P07490 rattus norv
3	61	32.8	1274	1	BXE_CLOBO	P30996 clostridium
4	61	32.8	1296	1	BXE_CLOBO	P60393 clostridium
5	59	31.7	1295	1	BXA1_CLOBO	P10845 clostridium
6	58.5	31.5	1250	1	BXE_CLOBO	P00496 clostridium
7	58	31.2	1290	1	BXB_CLOBO	P10844 clostridium
8	57.5	30.9	90	1	GONL_MOUSE	P13562 mus musculu
9	57	30.6	1051	1	VP2_AHSY6	O71024 african hor
10	56.5	30.4	1250	1	BXE_CLOBO	P30995 clostridium
11	56	30.1	1295	1	BXA2_CLOBO	P45894 clostridium
12	54.5	29.3	92	1	GONL_HUMAN	P01148 homo sapien
13	54.5	29.3	1290	1	BXC1_CLOBO	P18640 clostridium
14	54	29.0	67	1	GONL_MACMU	P55247 macaca mula
15	54	29.0	91	1	GONL_PIG	P49921 sus scrofa
16	53.5	28.8	760	1	ANV_CLOAB	P23671 clostridium
17	53.5	28.8	3122	1	DPOZ_MOUSE	P61493 mus musculu
18	52	28.0	61	1	GONL_SHEEP	O09163 mesocricetu
19	52	28.0	63	1	GONL_MESAU	P45656 xenopus lae
20	52	28.0	89	1	GONL_XENLA	P45656 xenopus lae
21	52	28.0	92	1	GONL_TUPGB	P45656 xenopus lae
22	52	28.0	449	1	VNSS_INSVN	P45656 xenopus lae
23	52	28.0	464	1	VNSS_TSWV1	P45656 xenopus lae
24	52	28.0	467	1	VNSS_TSWV1	P45656 xenopus lae
25	51.5	27.7	90	1	GON8_RANDY	P26002 tomato spot
26	51.5	27.7	521	1	UBP3_HUMAN	P26003 tomato spot
27	51	27.4	95	1	GONL_MORSA	O91614 homo sapien
28	51	27.4	292	1	CC22_ORYSA	O73812 morone saxa
29	51	27.4	1196	1	BXCN_CLOBO	P29619 oryza sativ
30	50.5	27.2	537	1	CORL_SCHPO	P46081 clostridium
31	50	26.9	241	1	CRTA_RHOCA	P07657 schizosacch
32	50	26.9	842	1	ABC6_HUMAN	P17055 rhodobacter
33	49.5	26.6	95	1	GON1_PAGMA	O9np58 homo sapien
						P70074 pagrus majo

34	49.5	26.6	95	1	GONL_SPAAU	P51919 sparus aura
35	49.5	26.6	3130	1	DPOZ_HUMAN	O60673 homo sapien
36	49	26.3	523	1	C9B1_GLYEC	P93149 glycyrrhiza
37	48.5	26.1	92	1	GONL_CAVPO	O54713 cavia porce
38	48.5	26.1	5217	1	HTS1_COCCA	Q01886 cochliobolu
39	48	25.8	10	1	GONL_ALLMI	P37041 alligator m
40	48	25.8	92	1	GONL_CHICK	P37042 gallus gall
41	48	25.8	94	1	GONL_HAPBU	P51918 haplochromi
42	48	25.8	99	1	GONL_DICLA	Q91a10 dicentrarch
43	48	25.8	280	1	CC2B_ANTMA	Q38773 antirrhinum
44	48	25.8	301	1	CRK1_LEIME	Q06309 leishmania
45	48	25.8	932	1	IF3A_SCHPO	O74760 schizosacch

ALIGNMENTS

RESULT 1
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).
OS Clostridium tetani.
OG Plasmid.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
RL EMOB J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CM3911;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE OF 742-1314 FROM N.A.
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RL J. Bacteriol. 165:21-27(1986).
RN [4]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Krieglstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";
RL Eur. J. Biochem. 188:39-45(1990).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
RT identification of cleavage sites.";
RL Eur. J. Biochem. 202:41-51(1991).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396558;
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
RT neurotransmitter release and protease activity depend on zinc.";

A:Accession: JH0256
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-27,'E',29-1251 <POU>
A:Cross-references: EMBL:X62088; NID:g40379
A:Experimental source: strains ATCC 43181 and ATCC 43755
R:Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa, N.
J. Gen. Microbiol. 137, 519-525, 1991
A:Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxin
A:Reference number: S16145; MUID:91237316
A:Accession: S16145
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229,'M',231-252 <FUG>
A:Cross-references: EMBL:X53180; NID:g40407; PIDN:CAA37321.1; PID:g40408
A:Experimental source: strain BL6340
A:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release
C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin
F:2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
F:423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>
F:412-426/Disulfide bonds: #status predicted

Query Match 30.4%; Score 56.5; DB 2; Length 1251;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;
QY 1 FNNFTVSFWLRVP-----KVSASHLE-----GP 23
: ||::|||::||
Db 912 YKNFSISFWRIPNYDNKIVNVNNEYTIINCMRDNSGKVSLSNHNHETIWTLODNGSINQ 971
: ||::|||::||
QY 24 SLHWSYG 30
: :||
Db 972 KLAIFYG 978

RESULT 15
S21178
botulinum neurotoxin type E precursor - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 30-Sep-1993 #sequence-revision 30-Sep-1993 #text-change 15-Oct-1999
C:Accession: S21178; S48107; JH0257; B35294; A60027; S18111
R:Wheilan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.
Eur. J. Biochem. 204, 657-667, 1992
A:Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin
A:Reference number: S21178; MUID:92174922
A:Accession: S21178
A:Molecule type: DNA
A:Residues: 1-1252 <WHE>
A:Cross-references: EMBL:X62683; NID:g40397; PIDN:CAA44558.1; PID:g40398
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin' gene and specific id
A:Reference number: S48103; MUID:94013372
A:Accession: S48107
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 616-982 <CAM>
A:Cross-references: EMBL:X70815; NID:g407786; PIDN:CAA50146.1; PID:g407787
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type
A:Reference number: JH0256; MUID:92181428
A:Accession: JH0257
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-176,'R',178-197,'C',199-339,'R',341-772,'I',774-962,'FE',965-966,'R',968-1
A:Cross-references: EMBL:X62089; NID:g40393; PIDN:CAA43999.1; PID:g40394
A:Experimental source: strain Beluga
R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990

A:Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
A:Reference number: A35294; MUID:90264400
A:Accession: B35294
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-176,'R',178-252 <BIN>
A:Experimental source: strain Beluga
R:Gimenez, J.A.; DasGupta, B.R.
Biochimie 72, 213-217, 1990
A:Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the
A:Reference number: A60027; MUID:90344918
A:Accession: A60027
A:Molecule type: protein
A:Residues: 420-427 <GIM>
A:Experimental source: strain Beluga
A:Note: this fragment was generated by proteolysis with Lys-C rather than with trypsi
C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit
C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin
F:1-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>
F:423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>
F:412-426/Disulfide bonds: #status predicted

Query Match 30.4%; Score 56.5; DB 2; Length 1252;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

QY 1 FNNFTVSFWLRVP-----KVSASHLE-----GP 23
: ||::|||::||
Db 912 YKNFSISFWRIPNYDNKIVNVNNEYTIINCMRDNSGKVSLSNHNHETIWTLODNGSINQ 971
: ||::|||::||
QY 24 SLHWSYG 30
: :||
Db 972 KLAIFYG 978

Search completed: October 10, 2002, 16:12:04
Job time : 9.99573 secs

A;Residues: 1:1291 <RES>
A;Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
A;Reference number: S48103; MUID:94013372
A;Accession: S48103
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 634-761, 'E', 763-841, 'M', 843, 'T', 845, 'N', 847-994 <CAM1>
A;Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779
A;Experimental source: non-proteolytic strain 2129B (Scott)
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A;Accession: S48104
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 634-843, 'T', 845, 'N', 847-994 <CAM2>
A;Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
A;Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)
C;Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic s
C;Genetics:
A;Gene: bont/b
C;Superfamily: tetanus toxin
C;Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>
F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVT>
F:230,234/Binding site: zinc (His) #status predicted
F:231/Active site: Glu #status predicted

Query Match 31.2%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 8.1;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
| :|:|||||:|:|
Db 923 FLDFSVSEWIRIPK 936

RESULT 11
RHWSG
gonadolibarin precursor - mouse
N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing
N;Contents: gonadolibarin; gonadolibarin-associated protein (GAP)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C;Accession: A47578
R;Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolics,
Science 234, 1366-1371, 1986
A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo
A;Reference number: A47578; MUID:87069928
A;Accession: A47578
A;Molecule type: DNA
A;Residues: 1-90 <MAS>
A;Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
C;Genetics:
A;Introns: 45/3; 77/3
C;Function:
A;Note: gonadolibarin stimulates pituitary secretion of lutropin and follitropin
C;Superfamily: gonadolibarin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadolibarin #status predicted <GLB>
F:35-90/Product: gonadolibarin-associated protein #status predicted <GAP>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 30.9%; Score 57.5; DB 1; Length 90;
Best Local Similarity 80.0%; Pred. No. 0.5;
Matches 12; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 20 LEG-PSLHWSYGLRP 33
| | | | | | | | | | | | | | | |

Db 16 LEGC3SQHWSYGLRP 30

RESULT 12
T36589
probable transmembrane protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
R;Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A;Reference number: 221575
A;Accession: T36589
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-502 <OLI>
A;Cross-references: EMBL:AL049826; PIDN:CAB42730.1; GSPDB:GN00070; SCOEDB:SCH24.24
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCH24.24
C;Superfamily: Mycobacterium tuberculosis probable membrane protein Rv0051

Query Match 30.6%; Score 57; DB 2; Length 502;
Best Local Similarity 50.0%; Pred. No. 3.9;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 18 SHLEGPSLHWSYGLRP 33
:| | | | | | | | | | | | | | | |
Db 483 AHFEGPQVHWGKGROP 498

RESULT 13
S48106
neurotoxin type E - Clostridium botulinum (fragment)
C;Species: Clostridium botulinum
C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C;Accession: S48106
R;Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A;Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A;Reference number: S48103; MUID:94013372
A;Accession: S48106
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-367 <CAM>
A;Cross-references: EMBL:X70818; NID:g407784; PIDN:CAA50149.1; PID:g407785
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

Query Match 30.4%; Score 56.5; DB 2; Length 367;
Best Local Similarity 22.4%; Pred. No. 3.3;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

QY 1 FNNFTVSFWLRVP-----KVSASHLE-----GP 23
| | | | | | | | | | | | | | | |
Db 297 YKNFSISFWIRIPNDKIVNNVEYTIINCMRDNSGKVSLSLHNHNEIWTQDNAGINQ 356
| :|:| | | | | | | | | | | | | | | |

QY 24 SLHWSYG 30
| :|:| | | | | | | | | | | | | | | |
Db 357 KLAFNYG 363

RESULT 14
JH0256
botulinum neurotoxin type E precursor - Clostridium butyricum
C;Species: Clostridium butyricum
C;Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
R;Foulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
C;Accession: JH0256; S16145
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A;Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum t
A;Reference number: JH0256; MUID:92181428

A:Molecule type: protein
A:Residues: 2-6:445-453,'X',455-457 <DASI>
R:DasGupta, B.R.; Foley, J.; Niece, R.
Biochemistry 26, 4162, 1987
A:Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A:Reference number: A27000
A:Accession: A27000
A:Molecule type: protein
A:Residues: 2-47 <DAS2>
R:Binz, T.; Blas, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
J. Biol. Chem. 269, 1617-1620, 1994
A:Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.
A:Reference number: A49708; MUID:94124495
A:Contents: annotation
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses
C:Genetics:
A:Gene: atx; botA
C:Function:
A:Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associated
C:Superfamily: tetanus toxin
C:Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane protein
F:2-44/Product: bontoxilysin A light chain #status experimental <LGHT>
F:445-1296/Product: bontoxilysin A heavy chain #status experimental <HVY>
F:223-227/Binding site: zinc (His) #status predicted
F:224/Active site: Glu #status predicted
Query Match 31.7% Score 59; DB 1; Length 1296;
Best Local Similarity 57.1%; Pred. No. 5.9;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 FNNFTVSFWLRVPK 14
: ||: |||: |||
Db 938 YENFTSFWRIPK 951
RESULT 9
A48940
bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
N:Alternate names: botulinum neurotoxin type B (BoNT/B)
C:Species: Clostridium botulinum
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A:Title: Molecular cloning of the Clostridium botulinum structural gene encoding the type
A:Reference number: A48940; MUID:92384550
A:Accession: A48940
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1291 <WHE>
A:Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
A:Experimental source: type B, Danish
A:Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIIP:112081); this publica
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
A:Reference number: S48105
A:Accession: S48105
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 634-994 <CAMS>
A:Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
A:Experimental source: proteolytic type B, strain NCTC 7273
R:Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A:Description: Partial amino acid sequence of botulinum neurotoxin type B and comparis
A:Reference number: S21575
A:Accession: S21575
A:Molecule type: DNA
A:Residues: 36-217,'G',219-224,'S',226-246 <SZA>
A:Cross-references: EMBL:211934; NID:g40383; PIDN:CAA77991.1; PID:g40384
R:Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K
J. Biol. Chem. 267, 14721-14729, 1992

A:Title: Minimal essential domains specifying toxicity of the light chains of tetanus
A:Reference number: A42871; MUID:92340509
A:Accession: A42871
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-313,'S',315-451 <KUR>
A:Note: sequence extracted from NCBI backbone (NCBIIP:109365)
R:DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A:Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity wi
A:Reference number: S07155; MUID:89000987
A:Accession: S07155
A:Molecule type: protein
A:Residues: 2-29,'M',31-45 <DAS>
A:Accession: S08562
A:Molecule type: protein
A:Residues: 442-463,'R',465-467 <DA2>
R:Schmidt, J.J.; Sathymoorthy, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A:Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A:Reference number: S07128; MUID:85197963
A:Accession: S07128
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-16 <SCH1>
A:Accession: S08573
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-17 <SCH2>
A:Accession: S08574
A:Status: preliminary
A:Molecule type: protein
A:Residues: 442-459 <SCH3>
R:Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,
Nature 359, 832-835, 1992
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A:Reference number: S27125; MUID:93063293
A:Contents: annotation
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy
C:Genetics:
A:Gene: bont/b
C:Function:
A:Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C:Superfamily: tetanus toxin
C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F:2-441/Product: bontoxilysin B light chain #status experimental <LGHT>
F:442-1291/Product: bontoxilysin B heavy chain #status experimental <HVY>
F:230-234/Binding site: zinc (His) #status predicted
F:231/Active site: Glu #status predicted
Query Match 31.2% Score 58; DB 1; Length 1291;
Best Local Similarity 64.3%; Pred. No. 8.1;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNNFTVSFWLRVPK 14
: ||: |||: |||
Db 923 FLDFSVSFWRIPK 936
RESULT 10
I40631
non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
A:Accession: I40631; S48103; S48104; S36015
R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Curr. Microbiol. 28, 101-110, 1994
A:Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botul
A:Reference number: I40631; MUID:94122659
A:Accession: I40631
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

S48109
 neurotoxin type F - Clostridium botulinum (fragment)
 C:Species: Clostridium botulinum
 C>Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
 C:Accession: S48109
 R:Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulin neurotoxin gene and specific id
 A:Reference number: S48103; MUID:94013372
 A:Accession: S48109
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-369 <CAM>
 A:Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 C:Superfamily: tetanus toxin

Query Match 32.8%; Score 61; DB 2; Length 369;
 Best Local Similarity 57.1%; Pred. No. 0.76;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
 :||:||||:|
 Db 297 YQNFISFWVRPK 310

RESULT 6
 I40813
 neurotoxin type F - Clostridium botulinum
 C:Species: Clostridium botulinum
 C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
 C:Accession: I40813; S48108
 R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.F.
 FEMS Microbiol. Lett. 96, 225-230, 1992
 A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
 A:Reference number: I40644
 A:Accession: I40813
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1274 <RES>
 A:Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
 R:Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulin neurotoxin gene and specific id
 A:Reference number: S48103; MUID:94013372
 A:Accession: S48108
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 634-1002 <CAM>
 A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 32.8%; Score 61; DB 2; Length 1274;
 Best Local Similarity 57.1%; Pred. No. 3;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPK 14
 :||:||||:|
 Db 930 YQNFISFWVRPK 943

RESULT 7
 S39791
 neurotoxin - Clostridium botulinum
 C:Species: Clostridium botulinum
 C>Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
 C:Accession: S39791
 R:Campbell, K.; Collins, M.D.; East, A.K.
 Biochim. Biophys. Acta 1216, 487-491, 1993
 A:Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridium a
 A:Reference number: S39791; MUID:94092745
 A:Accession: S39791

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1297 <CAM>
 A:Cross-references: EMBL:X74162; NID:g441275; PIDN:CAA52275.1; PID:g441276
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 32.8%; Score 61; DB 2; Length 1297;
 Best Local Similarity 38.1%; Pred. No. 3.1;
 Matches 8; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
 :||:||||:|
 Db 930 FDNFSINFWRTPKYNNNDIQ 950

RESULT 8
 BTCLAB
 bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum
 N:Alternate names: botulinum neurotoxin type A
 C:Species: Clostridium botulinum
 C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
 C:Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000
 R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
 J. Biol. Chem. 265, 9153-9158, 1990
 A:Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
 A:Reference number: A35294; MUID:90264400
 A:Accession: A35294
 A:Molecule type: DNA
 A:Residues: 1-1296 <BIN>
 A:Cross-references: GB:M30196; NID:g144864; PIDN:AAA23262.1; PID:g144865
 A:Experimental source: strain 62A, subtype A
 R:Thompson, D.E.; Brehm, J.K.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson,
 Eur. J. Biochem. 189, 73-81, 1990
 A:Title: The complete amino acid sequence of the Clostridium botulinum type A neuroto
 A:Reference number: S09492; MUID:90235864
 A:Accession: S09492
 A:Molecule type: DNA
 A:Residues: 1, 'Q', 3-26, 'V', 28-1296 <THO>
 A:Cross-references: EMBL:X52056; NID:g40381; PIDN:CAA36289.1; PID:g40382
 A:Experimental source: NCTC 2916
 R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
 FEBS Lett. 376, 41-44, 1995
 A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compone
 A:Reference number: S67988; MUID:96096783
 A:Accession: S68220
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-12 <FUJ>
 A:Cross-references: EMBL:D67030; DDBJ:D50421; NID:g2160224
 R:Betley, M.J.; Somers, E.; Dasgupta, B.R.
 Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
 A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-t
 A:Reference number: A33401; MUID:89350959
 A:Accession: A33401
 A:Molecule type: DNA
 A:Residues: 1-35 <BET>
 A:Cross-references: GB:M27892; NID:g144880; PIDN:AAA23269.1; PID:g551776
 R:Gimenez, J.A.; Dasgupta, B.R.
 J. Protein Chem. 12, 351-363, 1993
 A:Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,
 A:Reference number: A53884; MUID:94000342
 A:Accession: A53884
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 867-880; 1148-1217, 'Y', 1219 <GIM>
 A:Experimental source: strain Hall
 R:Dasgupta, B.R.; Dekleva, M.L.
 Biochimie 72, 661-664, 1990
 A:Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and a
 A:Reference number: A60025; MUID:91120847
 A:Accession: A60025

R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A:Reference number: S69348; MUID:95262688
A:Accession: S69348
A:Molecule type: protein
A:Residues: 2-31 <DEF>
C:Comment: The source of this protein was an extrachromosomal plasmid.
C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (fragment A) and heavy (fragment B) chains. The amino end of the heavy chain (fragment B) contains a zinc-binding site. Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglionic presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of synaptobrevin.
C:Function: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synaptobrevin.
C:Superfamily: tetanus toxin
C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F:2-457/Product: tentoxylisin light chain (fragment A) #status predicted <TTL>
F:461-864/Domain: channel forming (fragment B) #status experimental <TXB>
F:865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
F:233,237/Binding site: zinc (His) #status predicted
F:234/Active site: Glu #status predicted

Query Match 60.2%; Score 112; DB 1; Length 1315;
Best Local Similarity 100.0%; Pred. No. 1.9e-07; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
Db 947 FNNFTVSFWLRVPKVSASHLE 967
|||||

RESULT 2
botulinum neurotoxin type F - Clostridium baratii
C:Species: Clostridium baratii
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C:Accession: S33411; S31860
R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T. FEMS Microbiol. Lett. 108, 175-182, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium baratii type F neurotoxin
A:Reference number: S33411; MUID:93252228
A:Accession: S33411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1268 <THO>
A:Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 33.3%; Score 62; DB 2; Length 1268;
Best Local Similarity 64.3%; Pred. No. 2.2;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
:|||||:|
Db 922 YONFSISFWVRIPK 935
:|||||:|

RESULT 3
RHRTG
gonadoliberin precursor - rat
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormone N-terminus; gonadoliberin; prolactin release-inhibiting factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A40147; B26173; A48410
R:Bond, C.F.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P. Mol. Endocrinol. 3, 1257-1262, 1989
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic expression
A:Reference number: A40147; MUID:89384661
A:Accession: A40147

A:Molecule type: DNA
A:Residues: 1-92 <EON>
A:Cross-references: GB:M31670; NID:g204447; PIDN:AAA41264.1; PID:g204448
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H. Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor
A:Reference number: A94090; MUID:86094338
A:Accession: B26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADN>
A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E. Cell. Mol. Neurobiol. 12, 447-454, 1992
A:Title: Thyrocytes express a mRNA that is identical to hypothalamic lutelizing hormone
A:Reference number: A48410; MUID:93105480
A:Accession: A48410
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <MAI>
A:Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
A:Experimental source: thymus
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)
C:Genetics:
A:Introns: 47/3; 79/3
C:Function:
A:Description: stimulates pituitary secretion of lutropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: gonadoliberin #status predicted <PGN>
F:24-33/Product: gonadoliberin #status predicted <GLN>
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 33.1%; Score 61.5; DB 1; Length 92;
Best Local Similarity 51.7%; Pred. No. 0.14;
Matches 15; Conservative 3; Mismatches 4; Indels 7; Gaps 2;

Qy 12 VPKVSAS-----HLEG-PSLHWSYGLRP 33
:|||||:|
Db 4 IPKMAAVVLLTVCLGCSQHSYGLRP 32
:|||||:|

RESULT 4
S48110
neurotoxin type F - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C:Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S48110
R:Campbell, K.D.; Collins, M.D.; East, A.K. J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48110
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-366 <CAM>
A:Cross-references: EMBL:X70821; NID:g407792; PIDN:CAA50152.1; PID:g407793
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 32.8%; Score 61; DB 2; Length 366;
Best Local Similarity 57.1%; Pred. No. 0.76;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
:|||||:|
Db 297 YONFSISFWVRIPK 310
:|||||:|

RESULT 5

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 7.9188 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNFTVFWLRVFKVSASHLEGPGLHSYGLRXPX 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	60.2	1315	1 BTCLTN	tentoxylisin (EC 3
2	62	33.3	1268	2 S33411	botulinum neurotox
3	61.5	33.1	92	1 RHRTG	gonadolibarin prec
4	61	32.8	366	2 S48110	neurotoxin type F
5	61	32.8	369	2 S48109	neurotoxin type F
6	61	32.8	1274	2 I40813	neurotoxin type F
7	61	32.8	1297	2 S39791	neurotoxin - Clost
8	59	31.7	1296	1 BTCLAB	bontoxylisin (EC 3
9	58	31.2	1291	1 A48940	bontoxylisin (EC 3
10	58	31.2	1291	2 I40631	non-proteolytic bo
11	57.5	30.9	90	1 RHMSG	gonadolibarin prec
12	57	30.6	502	2 T36589	probable transmemb
13	56.5	30.4	367	2 S48106	neurotoxin type E
14	56.5	30.4	1251	2 JH0256	botulinum neurotox
15	56.5	30.4	1252	2 S21178	botulinum neurotox
16	56	30.1	812	2 T01618	hypothetical prote
17	56	30.1	1296	2 I40645	botulinum neurotox
18	54.5	29.3	92	1 RHUOG	gonadolibarin prec
19	54.5	29.3	1285	2 T70582	botulinum neurotox
20	54.5	29.3	1291	2 S46431	botulinum neurotox
21	54.5	29.3	1291	2 A49777	botulinum neurotox
22	54	29.0	67	2 I78541	gonadolibarin prec
23	53.5	28.8	469	2 B37837	probable alpha-amy
24	53.5	28.8	3122	2 T17202	DNA-directed DNA p
25	53	28.5	519	2 S78196	probable maturase
26	52	28.0	10	1 RHPGG	gonadolibarin - pi
27	52	28.0	10	1 RHSHG	gonadolibarin - sh
28	52	28.0	89	2 I51423	gonadolibarin prec
29	52	28.0	449	2 S23158	nucleocapsid prote

ALIGNMENTS

RESULT 1

BTCLTN

tentoxylisin (EC 3.4.24.68) precursor - Clostridium tetani

N:Alternate names: tetanus neurotoxin

C:Species: Clostridium tetani

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999

C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364

R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,

EMBO J. 5, 2495-2502, 1986

A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b

A:Reference number: A25689; MUID:87053814

A:Accession: A25689

A:Molecule type: DNA

A:Residues: 1-1315 <EIS>

A:Cross-references: GB:X04336; NID:g40769; PIDN:CAA28033.1; PID:g40770

R:Fairweather, N.F.; Lyness, V.A.

Nucleic Acids Res. 14, 7809-7812, 1986

A:Title: The complete nucleotide sequence of tetanus toxin.

A:Reference number: A25757; MUID:87040747

A:Accession: A25757

A:Molecule type: DNA

A:Residues: 1-1315 <FAI>

A:Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774

R:Experimental source: strain CN3911

R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.

J. Bacteriol. 165, 21-27, 1986

A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C 1

A:Reference number: A25194; MUID:86085672

A:Accession: A25194

A:Molecule type: DNA

A:Residues: 743-1315 <FA2>

A:Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921

A:Accession: B25194

A:Molecule type: protein

A:Residues: 865-894 <FA3>

R:Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.

Infect. Immun. 57, 3588-3593, 1989

A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termin

A:Reference number: A60759; MUID:90035436

A:Accession: A60759

A:Molecule type: protein

A:Residues: 461-475 <MAT>

R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.

J. Immunol. 142, 394-402, 1989

A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.

A:Reference number: J50098; MUID:89093918

A:Contents: annotation; epitope region

R:Schiavo, G.; Benfenati, F.; Poullain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta,

Nature 359, 832-835, 1992

A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo

A:Reference number: S27125; MUID:93063293

A:Contents: annotation


```
Best Local Similarity 35.9%; Pred. No. 6.1;
Matches 34; Conservative 4; Mismatches 7; Indels 14; Gaps 2;

QY 7 SFWLVPKV-----SASHLEG-----PSLHWSYGL 31
   ||| :|: ||| ||| :||| |
Db 224 SFWSFLPRTVWFSSSAWHLESORLEKLGTLHWKNGV 262
   ||| :|: ||| ||| :||| |

RESULT 13
Q9X8T8
ID Q9X8T8 PRELIMINARY; PRT; 502 AA.
AC Q9X8T8;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE PUTATIVE TRANSMEMBRANE PROTEIN.
GN SH24.24.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL049826; CAB42730.1; -.
KW Transmembrane.
SQ SEQUENCE 502 AA; 54795 MW; C84F774C0A5AADBC CRC64;

Query Match 30.6%; Score 57; DB 2; Length 502;
Best Local Similarity 50.0%; Pred. No. 7.2;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 18 SHLEGPSLHWSYGLRP 33
   :| ||| :|| | |
Db 483 AHFEGPQVHWKGRGP 498
   :| ||| :|| | |

RESULT 14
Q90Y63
ID Q90Y63 PRELIMINARY; PRT; 90 AA.
AC Q90Y63;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE GONADOTROPIN-RELEASING HORMONE.
GN GnRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21102951; PubMed=11170016;
RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GnRH1 and GnRH2
```

```
RT precursors from bullfrog (Rana catesbeiana).";
RL J. Exp. Zool. 289:150-201(2001).
DR EMBL; AF188754; AAL05972.1; -.
SQ SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;

Query Match 30.4%; Score 56.5; DB 13; Length 90;
Best Local Similarity 46.7%; Pred. No. 1.2;
Matches 14; Conservative 4; Mismatches 9; Indels 3; Gaps 2;

QY 5 TVSFWLVPKVSASH-LEGPSLHWSYGLRP 33
   || | : : ||| : | ||| |||
Db 6 TVVLLLAIVILLSSHHHGO--HWSYGLRP 33
   || | : : ||| : | ||| |||

RESULT 15
Q45862
ID Q45862 PRELIMINARY; PRT; 367 AA.
AC Q45862;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE E (FRAGMENT).
GN BONT/E.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TYPE E, HAZEN 36208 (ATCC 9564);
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70815; CAA50146.1; -.
DR HSP; P10845; 3BFA.
KW Neurotoxin.
FT NON_TER 1 367
FT NON_TER 367 367
SQ SEQUENCE 367 AA; 42854 MW; 0810595B3A865570 CRC64;

Query Match 30.4%; Score 56.5; DB 2; Length 367;
Best Local Similarity 22.4%; Pred. No. 6;
Matches 15; Conservative 8; Mismatches 7; Indels 37; Gaps 2;

QY 1 FNNFTVSFWLVRP-----KVSASHLE-----Gp 23
   : ||| :||| :| |
Db 297 YKNFSISFWVRIPNYDNKIVNVNNEYTIINCMRDNNNGKVSLSLNHNEIWTLODNAGINQ 356
   : ||| :||| :| |

QY 24 SLHWSYG 30
   | : ||
Db 357 KLAFNYG 363

Search completed: October 10, 2002, 16:09:56
Job time : 15.2222 secs
```

Query Match	31.2%	Score 58;	DB 2;	Length 1291;
Best Local Similarity	64.3%;	Pred. No. 15;		
Matches	9; Conservative	4; Mismatches	1; Indels	0; Gaps
QY	1 FNNFTVSFWLRVPK 14			
	: : : : : :			
DB	923 FLDFSFSFWIRPK 936			
RESULT 11				
ID	Q933K0 PRELIMINARY;	PRT;	1291 AA.	
AC	Q933K0;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DE	TYPE B CRYPTIC NEUROTOXIN.			
OS	Clostridium botulinum.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=593, AND 588;			
RA	Kirma N., Ferreira J.L., Baumstark B.R.;			
RT	"Characterization of six type A strains of Clostridium botulinum that			
RT	contain type B toxin gene sequences.";			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF300466; AAL11499.1; -			
DR	EMBL; AF300465; AAL11498.1; -			
DR	Neurotoxin.			
SW	SEQUENCE 1291 AA; 150843 MW; 7AC1737B0FA5A151 CRC64;			
Query Match	31.2%;	Score 58;	DB 2;	Length 1291;
Best Local Similarity	64.3%;	Pred. No. 15;		
Matches	9; Conservative	4; Mismatches	1; Indels	0; Gaps
QY	1 FNNFTVSFWLRVPK 14			
	: : : : :			
DB	923 FLDFSFSFWIRPK 936			
RESULT 12				
ID	Q9XAV1 PRELIMINARY;	PRT;	430 AA.	
AC	Q9XAV1;			
DT	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
DE	ALKANE 1-MONOXYGENASE (EC 1.14.15.1).			
GN	ALKB.			
OS	Pseudomonas fluorescens.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas.			
OX	NCBI_TaxID=294;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CHAO;			
RA	MEDLINE=21128526; PubMed=11207749;			
RT	Smiths T.H.M., Roethlisberger M., Witholt B., Van Beilen J.B.;			
RT	"Molecular screening for alkane hydroxylase genes in Gram-negative and			
RT	Gram-positive bacteria.";			
RT	Environ. Microbiol. 1:307-317(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CHAO;			
RA	Smiths T.H.M.;			
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ009579; CAB51045.2; -			
DR	Monooxygenase; Oxidoreductase.			
SW	SEQUENCE 430 AA; 48337 MW; 04BE950980783E86 CRC64;			
Query Match	30.6%;	Score 57;	DB 2;	Length 430;

```

RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL: X70814; CAA50145.1; -.
DR HSSP: P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 42175 MW; 533EA98735CD98E1 CRC64;

Query Match 31.2%; Score 58; DB 2; Length 361;
Best Local Similarity 64.3%; Pred. No. 3.6;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVVK 14
| : : : : : : : : : : : : : : : : : :
Db 290 FLDFSVSWIRIPK 303

RESULT 6
Q45848
ID Q45848 PRELIMINARY; PRT; 361 AA.
AC Q45848;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMELrel. 15, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TYPE B, NON-PROTEOLYTIC EKLUND 2B (COLWORTH 229);
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL: X70819; CAA50150.1; -.
DR HSSP: P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 42131 MW; A2EOFFFC81F9533D CRC64;

Query Match 31.2%; Score 58; DB 2; Length 361;
Best Local Similarity 64.3%; Pred. No. 3.6;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVVK 14
| : : : : : : : : : : : : : : : : : :
Db 290 FLDFSVSWIRIPK 303

RESULT 7
Q9X708
ID Q9X708 PRELIMINARY; PRT; 441 AA.
AC Q9X708;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=99343691; PubMed=10413679;
RA Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
RA Schiavo G.;
RT "Functional characterisation of tetanus and botulinum neurotoxins
RT binding domains.";
RL J. Cell Sci. 112:2715-2724(1999).
DR EMBL: AJ242628; CAB43706.1; -.
DR HSSP: P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1
FT NON_TER 441
SQ SEQUENCE 441 AA; 52772 MW; 721D0B468E8C95A4 CRC64;

Query Match 31.2%; Score 58; DB 2; Length 441;
Best Local Similarity 64.3%; Pred. No. 4.5;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVVK 14
| : : : : : : : : : : : : : : : : : :
Db 81 FLDFSVSWIRIPK 94

RESULT 8
Q9ZAJ8
ID Q9ZAJ8 PRELIMINARY; PRT; 1291 AA.
AC Q9ZAJ8;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE BONT PROTEIN.
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 3281 (ATCC 43757);
RX MEDLINE=98440323; PubMed=97677110;
RA Santos-Buelga J., Collins M.D., East A.K.;
RT "Characterization of the genes encoding the Botulinum neurotoxin
RT complex in a strain of clostridium botulinum producing type B & F
RT neurotoxins.";
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL: Y13630; CAA73968.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptidse.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1291 AA; 150840 MW; E4D3B0E46AB2E735 CRC64;

Query Match 31.2%; Score 58; DB 2; Length 1291;
Best Local Similarity 64.3%; Pred. No. 15;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVVK 14
| : : : : : : : : : : : : : : : : : :
Db 923 FLDFSVSWIRIPK 936

RESULT 9
Q08077
ID Q08077 PRELIMINARY; PRT; 1291 AA.
AC Q08077;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (EC 3.4.24.-) (BONT/B).
GN BONT/B.

```



```

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE TETANUS TOXIN (FRAGMENT).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA Shumin Z., Dianliang L.;
RT "Cloning and sequence analysis of tetanus toxin gene.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389424; AAK72964.2; -.
FT NON_TER 1
FT NON_TER 1310 1310
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 60.2%; Score 112; DB 2; Length 1310;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | |
Db 948 FNNFTVSFWLRVPKVSASHLE 968

RESULT 3
Q45851 PRELIMINARY; PRT; 1268 AA.
AC Q45851;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEUROTOXIN TYPE F.
OS BONT /F.
GN Clostridium baratii.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1561;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9325228; PubMed=8486245;
RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA Richardson P.T.;
RT "Nucleotide sequence of the gene coding for Clostridium baratii type F
RT neurotoxin: Comparison with other clostridial neurotoxins.";
RL FEMS Microbiol. Lett. 108:175-182(1993).
DR EMBL; X68262; CAA48329.1; -.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 33.3%; Score 62; DB 2; Length 1268;
Best Local Similarity 64.3%; Pred. No. 4;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: | | | | | | | |
Db 922 YQNFSSFWVRIPK 935

RESULT 4
Q57236 PRELIMINARY; PRT; 1278 AA.
AC Q57236; Q45863;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1998 (TrEMBLrel. 01, Last sequence update)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE F (BONT/F PROTEIN).
GN BONT/F.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NCTC 10281;
RA Hutson R.A., Collins M.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 635-1000 FROM N.A.
RC STRAIN=NCTC 1028; PubMed=8408542;
RX MEDLINE=94013372;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-27 FROM N.A.
RC STRAIN=LANGELAND;
RX MEDLINE=98404102; PubMed=9732534;
RA East A.K., Bhandari M., Hielm S., Collins M.D.;
RT "Analysis of the botulinum neurotoxin type F gene clusters in
RT proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT baratii.";
RL Curr. Microbiol. 37:262-268(1998).
DR EMBL; X81714; CAA57358.1; -.
DR EMBL; L35496; AAA23210.1; -.
DR EMBL; X70821; CAA50152.1; -.
DR EMBL; X90064; CAA67512.1; -.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Neurotoxin.
SQ SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 32.8%; Score 61; DB 2; Length 1278;
Best Local Similarity 57.1%; Pred. No. 5.5;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPK 14
: | | | | | | | |
Db 931 YQNFSSFWVRIPK 944

RESULT 5
Q45846 PRELIMINARY; PRT; 361 AA.
AC Q45846;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TYPE B, NON-PROTEOLYTIC 2129B (SCOTT);

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:59:26 ; Search time 13.2222 Seconds
(without alignments)
444.844 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSWLRVPKVSASHLEGPLHWSYGLRXP 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvrius:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	60.2	451	2 Q9LA13	Q9LA13 clostridium
2	112	60.2	1310	2 Q93N27	Q93N27 clostridium
3	62	33.3	1268	2 Q45851	Q45851 clostridium
4	61	32.8	1278	2 Q57236	Q57236 clostridium
5	58	31.2	361	2 Q45846	Q45846 clostridium
6	58	31.2	361	2 Q45848	Q45848 clostridium
7	58	31.2	441	2 Q9X708	Q9X708 clostridium
8	58	31.2	1291	2 Q9ZAJ8	Q9ZAJ8 clostridium
9	58	31.2	1291	2 Q08077	Q08077 clostridium
10	58	31.2	1291	2 Q93G71	Q93G71 clostridium
11	58	31.2	1291	2 Q933K0	Q933K0 clostridium
12	57	30.6	430	2 Q9XAV1	Q9XAV1 pseudomonas
13	57	30.6	502	2 Q9X8T8	Q9X8T8 streptomyces
14	56.5	30.4	90	13 Q90Y63	Q90Y63 rana catesb
15	56.5	30.4	367	2 Q45862	Q45862 clostridium
16	56.5	30.4	367	2 Q45861	Q45861 clostridium

17	56.5	30.4	1255	2 Q9FAR6	Q9far6 clostridium
18	56	30.1	812	10 O64620	O64620 arabidopsis
19	56	30.1	1251	2 Q9K395	Q9K395 clostridium
20	56	30.1	1251	2 Q9K395	Q9K395 clostridium
21	55	29.6	1280	2 Q45894	Q45894 clostridium
22	55	29.6	1280	2 Q9ZAJ5	Q9ZAJ5 clostridium
23	54.5	29.3	1285	2 Q45967	Q45967 clostridium
24	54.5	29.3	1285	2 Q9LBR1	Q9LBR1 clostridium
25	54	29.0	91	13 Q9PRH0	Q9prh0 anguilla ja
26	53.5	28.8	760	2 Q9S429	Q9s429 clostridium
27	52	28.0	467	12 Q37367	Q37367 tomato spot
28	52	28.0	467	12 Q37369	Q37369 tomato spot
29	51.5	27.7	91	13 Q9DGC8	Q9dgc8 oryza lat
30	51.5	27.7	317	4 Q9Y2R8	Q9y2r8 homo sapien
31	51	27.4	130	2 Q93PY5	Q93py5 pseudomonas
32	51	27.4	220	10 Q9C592	Q9c592 arabidopsis
33	51	27.4	367	16 Q987W1	Q987w1 rhizobium l
34	51	27.4	467	12 Q88900	Q88900 tospovirus.
35	51	27.4	496	3 Q90U16	Q9uu16 schizosacch
36	51	27.4	503	10 Q948E8	Q948e8 oryza sativ
37	51	27.4	659	16 Q67656	Q67656 aquifex aeo
38	51	27.4	944	5 Q17411	Q17411 caenorhabdi
39	51	27.4	1196	2 Q9LBS8	Q9lbs8 clostridium
40	51	27.4	1196	2 Q9LBR2	Q9lbr2 clostridium
41	51	27.4	1196	2 Q53550	Q53550 clostridium
42	51	27.4	1196	2 Q45916	Q45916 clostridium
43	51	27.4	1196	2 Q93HT4	Q93ht4 clostridium
44	51	27.4	1196	9 Q9ZX77	Q9zx77 clostridium
45	51	27.4	1196	9 Q38197	Q38197 clostridium

ALIGNMENTS

RESULT 1

Q9LA13					
ID	Q9LA13	PRELIMINARY;	PRT;	451	AA.
AC	Q9LA13;				
DT	01-OCT-2000 (Tremblrel. 15, Created)				
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)				
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)				
DE	TETANUS TOXIN (FRAGMENT).				
OS	Clostridium tetani.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;				
OX	Clostridium.				
NCBI_TaxID=1513;					
RP	SEQUENCE FROM N.A.				
RC	STRAIN=20886;				
RA	He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;				
RT	"Fragment C of Tetanus Toxin";				
RL	Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; AF154828; AAF3267.1; -.				
DR	HSP; P04958; 1A8D.				
DR	InterPro; IPR001064; Crystallin.				
DR	PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.				
FT	NON_TER				
SQ	SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;				

Query Match 60.2%; Score 112; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 1e-07; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 1 FNNFTVSWLRVPKVSASHLE 21
|||||

Db 83 FNNFTVSWLRVPKVSASHLE 103

RESULT 2

Q93N27					
ID	Q93N27	PRELIMINARY;	PRT;	1310	AA.
AC	Q93N27;				
DT	01-DEC-2001 (Tremblrel. 19, Created)				

CC Note: This sequence was constructed from the wild type human PSM
 CC (AA926T9), which appears on pages 184-187 of the specification.

XX SQ Sequence 750 AA;
 Query Match 61.8%; Score 115; DB 21; Length 750;
 Best Local Similarity 81.5%; Pred. No. 2.9e-08;
 Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 FNNFTVSFWLRVPKVSASHLEGPLHW 27
 |||||
 Db 87 FNNFTVSFWLRVPKVSASHLEVELAHY 113

RESULT 15

ID AAB20149 standard; Protein: 109 AA.

XX AC AAB20149;

XX DT 30-APR-2001 (first entry)

XX DE Growth differentiation factor 8 AutoVac construct GDF-8 P30-2.

XX KW Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
 KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
 KW cardiant; human; mutant; mutein.

XX OS Chimeric - Homo sapiens.

OS Chimeric - Clostridium tetani.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Region 1..48

FT /note= "identical to residues 267-314 of human

FT GDF-8"

FT Region 49..69

FT /note= "tetanus toxoid P2 epitope"

FT Region 70..109

FT /note= "identical to residues 336-375 of human

FT GDF-8"

FT Misc-difference 73

FT /note= "Cys-73 may be substituted by Ser to avoid

FT disulfide bond formation"

FT Misc-difference 90..91

FT /note= "optionally replaced by Glu-Gly"

FT WO200105820-A2.

XX PN 25-JAN-2001.

XX PD 20-JUL-2000; 2000WO-DK00413.

XX PF 20-JUL-1999; 99DR-0001014.

XX PR 26-JUL-1999; 99US-0145275.

XX XX (MEBI-) M & E BIOTECH AS.

XX Halkier T, Mouritsen S, Klysner S;

XX WPI; 2001-112680/12.

XX Increasing the muscle mass of animals used in meat production by down

XX regulating growth differentiation factor 8 (GDF-8) activity in the

XX animal through induction of anti-GDF-8 antibody production -

XX Example 1; Page 101-102; 110pp; English.

XX The present sequence is that of AutoVac construct GDF-8 P30-2,

XX comprising the 109 C-terminal amino acid residues of human

XX growth differentiation factor 8 (GDF-8) in which residues 49-69 are

XX replaced by the promiscuous tetanus toxin T-cell epitope P30 (see

XX AAB20144). It is an object of the invention to produce a

CC recombinant therapeutic vaccine that is capable of effecting
 CC down-regulation of GDF-8 in order to increase the muscle growth
 CC rate of farm animals. The vaccines (see AAB20145-53) are capable
 CC of breaking autotolerance against autologous GDF-8. They comprise
 CC the C-terminal portion of human GDF-8 in which a portion of the
 CC native sequence is replaced by a T-cell epitope such as P30, with
 CC minimal disturbance of the authentic 3-dimensional structure of
 CC the protein. Nucleic acids encoding the GDF-8 variants can be used
 CC for genetic immunisation of the animals. Down-regulation of GDF-8
 CC activity can increase muscle mass by up to at least 45% in cattle,
 CC pigs and poultry used for meat production, reducing the need for
 CC antibiotic feed-additives. Anti-GDF8 vaccines can be used to
 CC treat human diseases such as cancer cachexia where muscle atrophy is
 CC pronounced and for patients suffering from acute and chronic heart
 CC failure.

XX SQ Sequence 109 AA;

Query Match 61.8%; Score 114.5; DB 22; Length 109;

Best Local Similarity 95.8%; Pred. No. 3.9e-09;

Matches 23; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy 1 FNNFTVSFWLRVPKVSASHLE-GP 23

|||||

Db 49 FNNFTVSFWLRVPKVSASHLEAGP 72

Search completed: October 10, 2002, 16:05:08

Job time : 19.8718 secs

RESULT 13
 AAY92646
 ID AAY92646 standard; Protein; 750 AA.
 XX
 AC AAY92646;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Mutant human prostate specific membrane antigen construct, hPSM10.3.
 XX
 KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 674..688
 FT /label= P2
 FT /note= "foreign epitope"
 FT Peptide 210..230
 FT /label= P30
 FT /note= "foreign epitope"
 XX
 FN WO200020027-A2.
 XX
 XX 13-APR-2000.
 XX
 XX 05-OCT-1999; 99WO-DK00525.
 XX
 XX 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX
 XX (MEBI-) M & E BIOTECH AS.
 XX
 XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 PI
 XX WPI; 2000-349917/30.
 DR
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 XX Example 1; Page -: 220pp; English.
 PS
 XX
 CC AAY92627-49 are mutant immunogenized human prostate specific membrane
 CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
 CC P30). The immunogenic analogues of PSM can be used in the claimed method
 CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
 CC antibody binding regions and cysteine residues involved in disulfide
 CC bonds are preserved in the immunogenized forms. The method is used for
 CC inducing immune responses against weakly immunogenic cell-associated
 CC peptide antigens (PA) such as those associated with cancers
 CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
 CC method comprises effecting simultaneous presentation by antigen producing
 CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
 CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the respective
 CC PA and including at least one foreign T helper epitope are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer
 CC when the PA is human PSM, FGF8b and Her2, respectively.
 CC Note: This sequence was constructed from the wild type human PSM
 CC (AAY92619), which appears on pages 184-187 of the specification.
 XX
 XX Sequence 750 AA;

Query Match 62.1%; Score 115.5; DB 21; Length 750;
 Best Local Similarity 71.9%; Pred. No. 2.5e-08;

Matches 23; Conservative 6; Mismatches 2; Indels 1; Gaps 1;
 Qy 1 FNNFTVSEWLRVPKVSASHLEGPSLHWSYGLR 32
 |||||
 Db 210 FNNFTVSEWLRVPKVSASHLE-PADYFAPGVK 240
 RESULT 14
 AAY92634
 ID AAY92634 standard; Protein; 750 AA.
 XX
 AC AAY92634;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Mutant human prostate specific membrane antigen construct, hPSM1.2.
 XX
 KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 24..38
 FT /label= P2
 FT /note= "foreign epitope"
 FT Peptide 87..107
 FT /label= P30
 FT /note= "foreign epitope"
 XX
 FN WO200020027-A2.
 XX
 XX 13-APR-2000.
 XX
 XX 05-OCT-1999; 99WO-DK00525.
 XX
 XX 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX
 XX (MEBI-) M & E BIOTECH AS.
 XX
 XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 PI
 XX WPI; 2000-349917/30.
 DR
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 XX Example 1; Page -: 220pp; English.
 PS
 XX
 CC AAY92627-49 are mutant immunogenized human prostate specific membrane
 CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
 CC P30). The immunogenic analogues of PSM can be used in the claimed method
 CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
 CC antibody binding regions and cysteine residues involved in disulfide
 CC bonds are preserved in the immunogenized forms. The method is used for
 CC inducing immune responses against weakly immunogenic cell-associated
 CC peptide antigens (PA) such as those associated with cancers
 CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
 CC method comprises effecting simultaneous presentation by antigen producing
 CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
 CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the respective
 CC PA and including at least one foreign T helper epitope are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer
 CC when the PA is human PSM, FGF8b and Her2, respectively.

```
OS Homo sapiens.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 153..173
FT /label= P30
FT /note= "foreign epitope"
FT Peptide 549..563
FT /label= P2
FT /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
XX peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX P30). The immunogenic analogues of PSM can be used in the claimed method
XX as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX antibody binding regions and cysteine residues involved in disulfide
XX bonds are preserved in the immunogenized forms. The method is used for
XX inducing immune responses against weakly immunogenic cell-associated
XX peptide antigens (PA) such as those associated with cancers
XX (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX method comprises effecting simultaneous presentation by antigen producing
XX cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX B-cell group derived from the cell-associated PA; and (2) at least 1
XX first T helper cell group which is foreign to the animal. Analogues of
XX human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX part of all known and predicted CTL and B-cell epitopes of the respective
XX PA and including at least one foreign T helper epitope are also claimed.
XX The method is used to treat prostate, prostate/breast or breast cancer
XX when the PA is human PSM, FGF8b and Her2, respectively.
XX Note: This sequence was constructed from the wild type human PSM
XX (AAY92619), which appears on pages 184-187 of the specification.
XX
XX Sequence 693 AA;
XX
XX Query Match 62.1%; Score 115.5; DB 21; Length 693;
XX Best Local Similarity 71.9%; Pred. No. 2.3e-08;
XX Matches 23; Conservative 6; Mismatches 2; Indels 1; Gaps 1;
XX
XX 1 FNNFTVSFWLRVPKVSASHLEGPDLHWSYGLR 32
XX |||||||||||||||||||1: : : : :
XX 153 FNNFTVSFWLRVPKVSASHLE-PADYFAPGVK 183
XX
XX RESULT 12
XX AAY92645
XX ID AAY92645 standard; Protein; 750 AA.
XX
XX AAY92645;
XX
XX 10-AUG-2000 (first entry)
```

```
XX
XX DE
XX
XX KW Prostate specific membrane antigen construct, hPSM8.3.
XX KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
XX KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX OS Homo sapiens.
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 606..620
XX FT /label= P2
XX FT /note= "foreign epitope"
XX FT Peptide 210..230
XX FT /label= P30
XX FT /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
XX peptide antigens for the treatment of breast and prostate cancer
XX
XX Example 1; Page -: 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
XX P30). The immunogenic analogues of PSM can be used in the claimed method
XX as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
XX antibody binding regions and cysteine residues involved in disulfide
XX bonds are preserved in the immunogenized forms. The method is used for
XX inducing immune responses against weakly immunogenic cell-associated
XX peptide antigens (PA) such as those associated with cancers
XX (self-proteins), e.g. human prostate specific membrane antigen (PSM),
XX heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
XX method comprises effecting simultaneous presentation by antigen producing
XX cells (APCs) of the animals immune system of: (1) at least 1 CTL
XX (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
XX B-cell group derived from the cell-associated PA; and (2) at least 1
XX first T helper cell group which is foreign to the animal. Analogues of
XX human PSM, human Her2 and human/murine FGF8b comprising a substantial
XX part of all known and predicted CTL and B-cell epitopes of the respective
XX PA and including at least one foreign T helper epitope are also claimed.
XX The method is used to treat prostate, prostate/breast or breast cancer
XX when the PA is human PSM, FGF8b and Her2, respectively.
XX Note: This sequence was constructed from the wild type human PSM
XX (AAY92619), which appears on pages 184-187 of the specification.
XX
XX Sequence 750 AA;
XX
XX Query Match 62.1%; Score 115.5; DB 21; Length 750;
XX Best Local Similarity 71.9%; Pred. No. 2.5e-08;
XX Matches 23; Conservative 6; Mismatches 2; Indels 1; Gaps 1;
XX
XX 1 FNNFTVSFWLRVPKVSASHLEGPDLHWSYGLR 32
XX |||||||||||||||||||1: : : : :
XX 210 FNNFTVSFWLRVPKVSASHLE-PADYFAPGVK 240
XX
XX QY
XX Db
```

```

PN W0200155181-A2.
XX
PD 02-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US02937.
XX
PR 31-JAN-2000; 2000US-0179213.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Lal AA, Xiao L, Zhou Z;
XX
XX WPI; 2001-514557/56.
DR N-PSDB; AAH47055.
XX
XX New recombinant multivalent protein comprising antigenic determinants
PT derived from more than one stage in a life cycle of Plasmodium vivax,
PT useful as a vaccine for treating, preventing and reducing malarial
PT infection -
XX
XX Claim 5; Page 48-55; 59pp; English.
XX
XX The invention relates to recombinant multivalent proteins (I) that
CC stimulate an immune response to Plasmodium vivax. (I) comprises antigenic
CC determinants, fragments or conservative substitutions, derived from more
CC than one stage in a life cycle of a Plasmodium vivax parasite. (I) is
CC useful as a vaccine for stimulating an immune response, specifically a
CC protective immune response that confers increased resistance to infection
CC by Plasmodium parasites, such as P. vivax. (I) is especially useful in
CC the treatment, prevention and reduction of malarial infection, as
CC research or diagnostic reagents for the detection of Plasmodium species
CC in a biological sample, and for conferring immunity against multiple
CC stages of the malarial parasite. The antibodies produced are useful for
CC the detection or measurement of antigenic epitopes derived from one or
CC more stages in a life cycle of a parasite, particularly P. vivax. The
CC vaccine comprising the recombinant proteins, is cost-effective, health-
CC promoting intervention for controlling, preventing or treating the
CC incidence of malaria. The present sequence represents the amino acid
CC sequence of the recombinant protein ViVac2p, a multivalent and
CC multistage vaccine against P. vivax.
XX
SQ Sequence 2028 AA;
Query Match 62.48; Score 116; DB 22; Length 2028;
Best Local Similarity 88.08; Pred. No. 6.6e-08;
Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNNFTVSEWLRVPKVSASHLEGP SL 25
Db 1062 FNNFTVSEWLRVPKVSASHLENSV 1086
RESULT 10
AA92647
ID AA92647 standard; Protein; 693 AA.
XX
AC AA92647;
XX
XX 10-AUG-2000 (first entry)
XX
XX Mutant human PSM antigen splice variant construct, hPSM'6.3.
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer; PSM;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH 153..173
FT Peptide /label= P30
FT /note= "foreign epitope"

```

```

FT Peptide 391..405
FT /label= P2
XX /note= "foreign epitope"
PN
XX W0200020027-A2.
XX
PD 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK00525.
XX
XX 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
PT
XX
XX Example 1; Page -; 220pp; English.
XX
XX AA92627-49 are mutant immunogenized human prostate specific membrane
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
CC P30). The immunogenic analogues of PSM can be used in the claimed method
CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
CC antibody binding regions and cysteine residues involved in disulfide
CC bonds are preserved in the immunogenized forms. The method is used for
CC inducing immune responses against weakly immunogenic cell-associated
CC peptide antigens (PA) such as those associated with cancers
CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
CC method comprises effecting simultaneous presentation by antigen producing
CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC B-cell group derived from the cell-associated PA; and (2) at least 1
CC first T helper cell group which is foreign to the animal. Analogues of
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC part of all known and predicted CTL and B-cell epitopes of the respective
CC PA and including at least one foreign T helper epitope are also claimed.
CC The method is used to treat prostate, prostate/breast or breast cancer.
CC When the PA is human PSM, FGF8b and Her2, respectively.
CC Note: This sequence was constructed from the wild type human PSM
CC (AA92619), which appears on pages 184-187 of the specification.
XX
SQ Sequence 693 AA;
Query Match 62.18; Score 115.5; DB 21; Length 693;
Best Local Similarity 71.9%; Pred. No. 2.3e-08;
Matches 23; Conservative 6; Mismatches 2; Indels 1; Gaps 1;
Qy 1 FNNFTVSEWLRVPKVSASHLEGP SLHWSYGLR 32
Db 153 FNNFTVSEWLRVPKVSASHLE-PADYFAPGVR 183
RESULT 11
AA92648
ID AA92648 standard; Protein; 693 AA.
XX
XX AA92648;
XX
XX 10-AUG-2000 (first entry)
XX
XX Mutant human PSM antigen splice variant construct, hPSM'8.3.
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer; PSM;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX

```

Synthetic.
Clostridium tetani.
Mus musculus.

Key Location/Qualifiers
Peptide 1..14
Protein /note= "His tag"
15..112
Peptide /note= "residues 158-255 of murine OPGL"
113..133
Protein /note= "tetraxin toxoid P30 epitope"
134..188
/note= "residues 262-316 of murine OPGL"

WO200015807-A1.
23-MAR-2000.
13-SEP-1999; 99WO-DK00481.
15-SEP-1998; 98DK-0001164.
02-OCT-1998; 98US-0102896.
(MEBT-) M & E BIOTECH AS.

Halkier T, Haaning J;
WPI: 2000-271444/23.
N-PSDB; AA299970.

In vivo down-regulation of osteoprotegerin ligand (OPGL) activity used to treat, prevent and ameliorate osteoporosis -
Example; Page 94-95; 110pp; English.

The present sequence represents fusion protein of murine osteoprotegerin ligand (OPGL) and tetanus toxoid P30 epitope. Osteoprotegerin is a secreted member of the tumour necrosis factor receptor family, which blocks osteoclastogenesis in a dose dependent manner. The OPGL protein is synthesised as a type II transmembrane protein. The murine and human CC OPGL polypeptides are 87% homologous. OPGL is a potent osteoclast CC differentiation factor when combined with CSF-1. It is not capable of CC inducing osteoclast differentiation in the absence of CSF-1. OPGL is CC also an activator of mature osteoclasts. The specification describes a CC method for the in vivo down-regulation of OPGL activity in an animal. CC The method comprises using at least one OPGL polypeptide or subsequence, CC and/or at least one OPGL analogue to induce an immune response in the CC animal. The method and OPGL polypeptide are useful for treating, CC preventing and ameliorating osteoporosis or other diseases or conditions CC characterised by excessive bone resorption.

Query Match 62.4%; Score 116; DB 21; Length 188;
Best Local Similarity 82.1%; Pred. No. 4.2e-09;
Matches 23; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

QY 1 FNNFTVSFWLRVPKVSASHLEGP SLHWS 28
|||||
Db 113 FNNFTVSFWLRVPKVSASHLE----NWS 136
:||

RESULT 8
AAB85697
ID AAB85697 standard; Protein; 1807 AA.
XX AC AAB85697;
XX 29-OCT-2001 (first entry)
DT DT
DE Recombinant protein ViVac1p.
KW Multivalent protein; immune response; Plasmodium vivax; parasite;
XX OS Plasmodium vivax.

protozoacide; vaccine; malaria; recombinant; ViVac1.
Synthetic.
Plasmodium vivax.
WO200155181-A2.
02-AUG-2001.
29-JAN-2001; 2001WO-US02937.
31-JAN-2000; 2000US-0179213.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
Lal AA, Xiao L, Zhou Z;
WPI: 2001-514557/56.
N-PSDB; AAH47058.

New recombinant multivalent protein comprising antigenic determinants derived from more than one stage in a life cycle of Plasmodium vivax, useful as a vaccine for treating, preventing and reducing malarial infection -
Claim 5; Page 39-45; 59pp; English.

The invention relates to recombinant multivalent proteins (I) that stimulate an immune response to Plasmodium vivax. (I) comprises antigenic determinants, fragments or conservative substitutions, derived from more than one stage in a life cycle of a Plasmodium vivax parasite. (I) is useful as a vaccine for stimulating an immune response, specifically a protective immune response that confers increased resistance to infection by Plasmodium parasites, such as P. vivax. (I) is especially useful in the treatment, prevention and reduction of malarial infection, as research or diagnostic reagents for the detection of Plasmodium species in a biological sample, and for conferring immunity against multiple for stages of the malarial parasite. The antibodies produced are useful for the detection or measurement of antigenic epitopes derived from one or more stages in a life cycle of a parasite, particularly P. vivax. The vaccine comprising the recombinant proteins, is cost-effective, health-promoting intervention for controlling, preventing or treating the incidence of malaria. The present sequence represents the amino acid sequence of the recombinant protein ViVac1p, a multivalent and multistage vaccine against P. vivax.

Query Match 62.4%; Score 116; DB 22; Length 1807;
Best Local Similarity 88.0%; Pred. No. 5.7e-08;
Matches 22; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLEGP SL 25
|||||
Db 841 FNNFTVSFWLRVPKVSASHLESV 865
|||||

RESULT 9
AAB85698
ID AAB85698 standard; Protein; 2028 AA.
XX AC AAB85698;
XX 29-OCT-2001 (first entry)
DT DT
DE Recombinant protein ViVac2p.
KW Multivalent protein; immune response; Plasmodium vivax; parasite;
XX OS Plasmodium vivax.

XX Multivalent protein; immune response; Plasmodium vivax; parasite; parasite;

XX 05-OCT-1999; 99WO-DK00525.
 PF 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX (MEBI-) M & E BIOTECH AS.
 XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX WPI; 2000-349917/30.
 DR
 XX
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 PS Example 1; Page 118; 220pp; English.
 XX
 CC AAY92650-55 are peptides designed which correspond to the P2 and P30
 CC epitopes with 5 flanking human prostate specific membrane antigen (hPSM)
 CC amino acids in each end. The flanking amino acids correspond to the
 CC epitope insertion sites 6, 8 and 10. The peptides will be used in, e.g.
 CC T cell proliferation assays, but also for ELISA or other in vitro
 CC assays. The claims detail a method for inducing immune responses against
 CC weakly immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, hPSM,
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
 CC method comprises effecting simultaneous presentation by antigen
 CC producing cells (APCs) of the animals immune system of: (1) at least 1
 CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the
 CC respective PA and including at least one foreign T helper epitope (e.g.
 CC P2 and/or P30) are also claimed. The method is used to treat prostate,
 CC prostate/breast or breast cancer when the PA is human PSM, FGF8b and
 CC Her2, respectively.
 XX
 SQ Sequence 31 AA;
 Query Match 63.7%; Score 118.5; DB 21; Length 31;
 Best Local Similarity 92.3%; Pred. No. 2.3e-10;
 Matches 24; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 FNNFTVSFWLRVPKVSASHLEPSLH 26
 DB 6 FNNFTVSFWLRVPKVSASHLE-PSH 30
 RESULT 4
 AAW81336
 ID AAW81336 standard; Protein; 158 AA.
 XX
 AC AAW81336;
 XX
 DT 21-APR-1999 (first entry)
 XX
 DE TNF30-5, a TNF-alpha analogue.
 XX
 KW Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
 KW vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
 KW cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
 KW asthma.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX WO9846642-A1.
 PN
 XX 22-OCT-1998.
 PD
 XX 15-APR-1998; 98WO-DK00157.
 PF

XX 24-APR-1997; 97US-0044187.
 PR 15-APR-1997; 97DK-0000418.
 XX (FERR) FARM LAB FERRING AS.
 XX Dalum I, Elsner H, Jensen MR, Mouritsen S;
 XX WPI; 1998-594561/50.
 DR N-PSDB; AAV68425.
 XX
 XX Modified human tumour necrosis factor-alpha - comprises
 PT immunodominant T cell epitope, useful in vaccines to treat or
 PT prevent TNF-associated diseases, e.g. cancer
 XX
 PS Claim 15; Page 81-82; 134pp; English.
 XX
 CC The present sequence represents a modified human tumour necrosis
 CC factor-alpha (TNF-alpha) analogue. The analogues have no residual
 CC TNF activity and are immunogenic in a large proportion of the human
 CC population (by using promiscuous epitopes). The TNF-alpha analogue
 CC is able to generate, in humans, neutralizing antibodies to wild-type
 CC human TNF alpha, has at least one fragment of TNF substituted by a
 CC peptide containing an immunodominant T-cell epitope, and at least one
 CC TNF-alpha B-cell epitope. The substitution causes a significant change
 CC in the amino acid sequence of any one of the strands in the front
 CC beta-sheet, any of the connecting loops or any of the B', I or D strands
 CC in the back beta-sheet. The TNF-alpha analogues are used as vaccines for
 CC treatment or prevention of diseases associated with excessive release
 CC or activity of TNF-alpha, e.g. rheumatoid arthritis, Crohn's disease,
 CC ulcerative colitis, cancer of any sort, disseminated sclerosis, diabetes,
 CC psoriasis, osteoporosis and asthma.
 XX
 SQ Sequence 158 AA;
 Query Match 63.4%; Score 118; DB 19; Length 158;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FNNFTVSFWLRVPKVSASHLEG 22
 DB 133 FNNFTVSFWLRVPKVSASHLEG 154
 RESULT 5
 AAY92627
 ID AAY92627 standard; Protein; 750 AA.
 XX
 AC AAY92627;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Mutant human prostate specific membrane antigen construct, hPSM1.1.
 XX
 KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Peptide 17..31 /label= p2
 FT /note= "foreign epitope"
 FT 32..52 /label= p30
 FT /note= "foreign epitope"
 XX
 XX WO200020027-A2.
 PN
 XX 13-APR-2000.
 PD
 XX

PA (ZAMB/) ZAMB T.
 XX Ladd AE, Wang CY, Zamb T;
 XX WPI; 1994-357910/44.
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -
 PT that suppress LHRH activity in males and females
 XX
 XX Claim 8; Page 84; 213pp; English.
 XX
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein hapten containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)
 CC which elicits an immune response to the coupled peptide in members of
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the
 CC invasive and Th domains and between the immune stimulator and hapten
 CC components. When the hapten is LHRH, then optionally the invasive domain
 CC can be omitted from the immune stimulator component.
 CC The present sequence represents an LHRH-containing, invasive-free
 CC immunogenic peptide as above which can be used as a potent vaccine for
 CC treating e.g. prostatic hyperplasia, androgen-dependent carcinoma,
 CC prostatic carcinoma, testicular carcinoma, endometriosis, benign
 CC uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility.
 XX
 SQ Sequence 32 AA;
 Query Match 81.7%; Score 152; DB 15; Length 32;
 Best Local Similarity 87.9%; Pred. No. 2.8e-15;
 Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 QY 1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33
 |||||
 DB 3 FNNFTVSFWLRVPKVSASHLE---HWSYGLRP 31
 RESULT 2
 AAY92633
 ID AAY92633 standard; Protein: 750 AA.
 XX
 AC AAY92633;
 XX
 XX 10-AUG-2000 (first entry)
 XX
 DE Mutant human prostate specific membrane antigen construct, hPSM1.10.
 XX
 KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX
 XX Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Peptide 24..38
 FT /label= P2
 FT /note= "foreign epitope"
 FT Peptide 673..693
 FT /label= P30
 FT /note= "foreign epitope"
 XX
 PN WO200020027-A2.
 XX
 PD 13-APR-2000.
 XX
 XX 05-OCT-1999; 99WO-DK00525.
 XX

PR 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 PA (MEBI-) M & E BIOTECH AS.
 XX
 XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX WPI; 2000-349917/30.
 XX
 XX Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 XX
 XX Example 1; Page -: 220pp; English.
 XX
 CC AAY92627-49 are mutant immunogenized human prostate specific membrane
 CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or
 CC P30). The immunogenic analogues of PSM can be used in the claimed method
 CC as an autovaccine to induce a CTL response. Subdominant CTL epitopes,
 CC antibody binding regions and cysteine residues involved in disulfide
 CC bonds are preserved in the immunogenized forms. The method is used for
 CC inducing immune responses against weakly immunogenic cell-associated
 CC peptide antigens (PA) such as those associated with cancers
 CC (self-proteins), e.g. human prostate specific membrane antigen (PSM),
 CC heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b). The
 CC method comprises effecting simultaneous presentation by antigen producing
 CC cells (APCs) of the animals immune system of: (1) at least 1 CTL
 CC (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
 CC B-cell group derived from the cell-associated PA; and (2) at least 1
 CC first T helper cell group which is foreign to the animal. Analogues of
 CC human PSM, human Her2 and human/murine FGF8b comprising a substantial
 CC part of all known and predicted CTL and B-cell epitopes of the respective
 CC PA and including at least one foreign T helper epitope are also claimed.
 CC The method is used to treat prostate, prostate/breast or breast cancer
 CC when the PA is human PSM, FGF8b and Her2, respectively.
 CC Note: This sequence was constructed from the wild type human PSM
 CC (AAY92619), which appears on pages 184-187 of the specification.
 XX
 SQ Sequence 750 AA;
 Query Match 65.3%; Score 121.5; DB 21; Length 750;
 Best Local Similarity 86.2%; Pred. No. 3.2e-09;
 Matches 25; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 1 FNNFTVSFWLRVPKVSASHLEGPSLHWSY 29
 |||||
 DB 673 FNNFTVSFWLRVPKVSASHLE-PSHNKY 700
 RESULT 3
 AAY92655
 ID AAY92655 standard; Peptide: 31 AA.
 XX
 AC AAY92655;
 XX
 XX 10-AUG-2000 (first entry)
 XX
 DE PSMpep012 - P30 inserted in hPSM insertion position 10.
 XX
 KW Foreign epitope; P2; prostate specific membrane antigen; vaccination;
 KW cytotoxic T-lymphocyte immunity; self-protein; cancer; breast cancer;
 KW prostate cancer; cell-associated peptide antigen.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Peptide 6..26
 FT /label= P30
 XX
 PN WO200020027-A2.
 XX
 XX 13-APR-2000.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:10 ; Search time 17.8718 Seconds
(without alignments)
211.311 Million cell updates/sec

Title: US-09-848-834A-10

Perfect score: 186

Sequence: 1 FNFETVFWLRVFKVNSASHLEGPSLHWSYGLRFX 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	81.7	32	15	AA62702
2	121.5	65.3	750	21	AA92633
3	118.5	63.7	31	21	AA92655
4	118	63.4	158	19	AAW81336
5	117	62.9	750	21	AA92627
6	117	62.9	750	21	AA92636
7	116	62.4	188	21	AA84423
8	116	62.4	1807	22	AA85697
9	116	62.4	2028	22	AA85698
10	115.5	62.1	693	21	AA92647
11	115.5	62.1	693	21	AA92648

12	115.5	62.1	750	21	AA92645	Mutant human prost
13	115.5	62.1	750	21	AA92646	Mutant human prost
14	115	61.8	750	21	AA92634	Mutant human prost
15	114.5	61.6	109	22	AA320149	Growth differentia
16	114	61.3	158	19	AAW81333	TNF30-2, a TNF- α p
17	114	61.3	173	21	AA926426	An osteoprotegerin
18	112	60.2	21	12	AA926426	Immunogenic conju
19	112	60.2	21	17	AAW06130	Tetanus toxoid pro
20	112	60.2	21	17	AAW06130	T-cell antigen TT3
21	112	60.2	21	18	AAW46449	Broad range helper
22	112	60.2	21	19	AAW67034	Tetanus toxin frag
23	112	60.2	21	20	AAW67579	T-cell epitope pep
24	112	60.2	21	20	AAW73222	Tetanus toxoid epi
25	112	60.2	21	21	AAW45512	Tetanus P30 epitop
26	112	60.2	21	21	AAW99876	Tetanus toxin T ce
27	112	60.2	21	21	AAW92626	Foreign epitope P3
28	112	60.2	21	21	AAW84428	Amino acid sequenc
29	112	60.2	21	21	AAW49260	CD4+ T cell epitop
30	112	60.2	21	22	AAE11764	Clostridium tetani
31	112	60.2	21	22	AAW85702	Amino acid sequenc
32	112	60.2	21	22	AAW85453	Universal tetanus
33	112	60.2	21	22	AAW85198	Tetanus Toxoid uni
34	112	60.2	21	22	AAW20144	Tetanus toxin T-ce
35	112	60.2	21	22	AAW86637	HER-2 B cell pepti
36	112	60.2	21	22	AAW46173	Tetanus toxoid TT9
37	112	60.2	21	22	AAW49072	Tetanus toxoid TT
38	112	60.2	28	22	AAW46176	Tetanus toxoid 947
39	112	60.2	31	21	AAW92653	PSMpep010 - P30 in
40	112	60.2	31	21	AAW92654	PSMpep011 - P30 in
41	112	60.2	33	22	AAW49075	Amyloid beta/tetan
42	112	60.2	36	22	AAW63662	Peptide comprising
43	112	60.2	36	22	AAW63515	A peptide which ma
44	112	60.2	43	22	AAW46177	Tetanus toxoid 830
45	112	60.2	43	22	AAW49076	Amyloid beta/tetan

ALIGNMENTS

RESULT 1
AA62702
ID AA62702 standard; peptide; 32 AA.
XX
AC AA62702;
XX
DT 10-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
KW Helper T cell epitope; universal immune stimulator; invasin; haptene;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility;
KW tetanus toxin.
XX
OS Synthetic.
XX
FH Key
FT Domain 1..22
FT Domain /note= "tetanus toxin helper T cell epitope"
FT Domain 23..32
FT Domain /note= "LHRH haptene"
XX
PN WO9425060-A.
XX
PD 10-NOV-1994.
XX
PF 28-APR-1994; 94WO-US04832.
XX
PR 27-APR-1993; 93US-0057166.
PR 14-APR-1994; 94US-0229275.
XX
PA (LADD/) LADD A E.
PA (WANG/) WANG C Y.

; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-618-312A-4

Query Match 60.2%; Score 112; DB 1; Length 452;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNETVSFWLRVPKVSASHLE 21
Db 84 FNNETVSFWLRVPKVSASHLE 104

Search completed: October 10, 2002, 16:14:01
Job time : 7.19231 secs

```
;
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-11703-64

Query Match 60.2%; Score 112; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
Db 7 FNNFTVSFWLRVPKVSASHLE 27

RESULT 14
US-07-618-312A-2
; Sequence 2, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
```

```
;
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-618-312A-2

Query Match 60.2%; Score 112; DB 1; Length 452;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
Db 84 FNNFTVSFWLRVPKVSASHLE 104

RESULT 15
US-07-618-312A-4
; Sequence 4, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,082
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/661,052
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION/DOCKET NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-09-188-082-12

Query Match 60.2%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 11
PCT-US93-11703-66
; Sequence 66, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant D. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-11703-66

Query Match 60.2%; Score 112; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 13
PCT-US93-11703-64
; Sequence 64, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
```



```

; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
US-09-382-855-5

Query Match 60.2%; Score 112; DB 4; Length 21.;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels

QY 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 9
US-09-183-714B-5
; Sequence 5, Application US/09183714B
; Patent No. 6221593
; GENERAL INFORMATION:
; APPLICANT: Boon-Falleur, Thierry
; APPLICANT: Brasseur, Francis
; APPLICANT: Rimoldi, Donata

```

```
%  
% African American men, assessed  
% TITLE OF INVENTION: Method for Determining Cancer by Determining Expression  
% TITLE OF INVENTION: of MAGE-10  
% FILE REFERENCE:
```

; CURRENT FILING DATE: 1998-10-30
 ;
 ; PRIOR APPLICATION NUMBER: US 08/724,774
 ; PRIOR FILING DATE: 1996-10-03

```

; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

Query Match	60.2%	Score 112;	DB 4;	Length 21;
Best Local Similarity	100.0%;	Pred. No. 3.6e-10;		
Matches 21;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qv	1	FNNFTVSEFLRVPKVSASHLE	21	

21 1 FNNETVSEFWLRVPKVSASHLE 21

05/07/88 002 12
: Application US/09188082
: Sequence 12, Patent No. 6270765
: GENERAL INFORMATION:
: APPLICANT: Yashwant M. Deo
: APPLICANT: Joel Goldstein
: APPLICANT: Robert Graziano
: APPLICANT: Chезian Somasundaram
: TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
: TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street, Suite 510
: CITY: Boston
: STATE: Massachusetts

```

;
;
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
;

```

```

;
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;

```

ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MXI-043CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-661-052-12

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 5
US-08-460-502-8
Sequence 8, Application US/08460502
Patent No. 5843464
GENERAL INFORMATION:
APPLICANT: Bakaletz, Lauren O.
APPLICANT: Kaumaya, Parvin T.
TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter and Griswold
STREET: 800 Superior Avenue
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2688

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,502
FILING DATE:
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Goirick, Mary E.
REGISTRATION NUMBER: 34,829
REFERENCE/DOCKET NUMBER: 22727/00120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8458
TELEFAX: (216) 241-0816
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-502-8

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 6
US-08-724-774B-5
Sequence 5, Application US/08724774B
Patent No. 5908778
GENERAL INFORMATION:
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie,
APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
APPLICANT: Stefan; Reed, Daryl
TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
TITLE OF INVENTION: Rejection Antigen precuros Mage-10,
TITLE OF INVENTION: Antibodies Specific To The Molecule, and
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,774B
FILING DATE: 03-October-1996
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5908778man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-3200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: amino acid
TOPOLOGY: linear

US-08-724-774B-5

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
| | | | | | | | | | | | | | | | | | | | | |
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 7
US-09-089-595-5
Sequence 5, Application US/09089595
Patent No. 6153728
GENERAL INFORMATION:
APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie,
APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
APPLICANT: Stefan; Reed, Daryl
TITLE OF INVENTION: MAGE-10 ENCODING cDNA, The Tumor
TITLE OF INVENTION: Rejection Antigen precuros Mage-10,
TITLE OF INVENTION: Antibodies Specific To The Molecule, and
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA

```
RESULT 2
US-08-488-351A-14
; Sequence 14, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-14
;
; Query Match 81.7%; Score 152; DB 2; Length 32;
; Best Local Similarity 87.9%; Pred. No. 7.7e-16;
; Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
;
Qy 1 FNNFTVFWLRVPKVSASHLEGPSLHWSYGLRP 33
Db 3 FNNFTVFWLRVPKVSASHLE----HWSYGLRP 31
;
RESULT 3
US-07-610-525-1
; Sequence 1, Application US/07610525
; Patent No. 5196512
; GENERAL INFORMATION:
; APPLICANT: BIANCHI Eisabetta
; APPLICANT: PESSI Antonello
; APPLICANT: CORRADIN Giampietro
; TITLE OF INVENTION: SYNTHETIC PEPTIDES USEFUL AS
```

```

; TITLE OF INVENTION: UNIVERSAL CARRIERS FOR THE PREPARATION OF IMMUNOGENIC CONJU
; TITLE OF INVENTION: AND THEIR USE IN THE DEVELOPMENT OF SYNTHETIC VACCINES.
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHEA & GOULD
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10020-1193
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/610,525
; FILING DATE: 19901108
; CLASSIFICATION: 424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-827-3000
; TELEFAX: 212-840-6702
; TELEX: 423973
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acid residues
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: no
; FRAGMENT TYPE: Internal fragment
US-07-610-525-1
;
; Query Match 60.2%; Score 112; DB 1; Length 21;
; Best Local Similarity 100.0%; Pred. No. 3.6e-10;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy 1 FNNFTVFWLRVPKVSASHLE 21
Db 1 FNNFTVFWLRVPKVSASHLE 21
;
RESULT 4
US-08-661-052-12
; Sequence 12, Application US/08661052
; Patent No. 5837243
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/661,052
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:46 ; Search time 7.19231 Seconds
(without alignments)
115.466 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSFWLRPKVSASHLEGPSLHWSYGLRXP 34

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	81.7	32	1	US-08-446-692-14
2	152	81.7	32	2	US-08-488-351A-14
3	112	60.2	21	1	US-07-610-525-1
4	112	60.2	21	2	US-08-661-052-12
5	112	60.2	21	2	US-08-460-502-8
6	112	60.2	21	2	US-08-724-774B-5
7	112	60.2	21	4	US-09-089-595-5
8	112	60.2	21	4	US-09-383-855-5
9	112	60.2	21	4	US-09-183-714B-5
10	112	60.2	21	4	US-09-188-082-12
11	112	60.2	21	4	US-09-171-969-10
12	112	60.2	21	5	PCT-US93-11703-66
13	112	60.2	31	5	PCT-US93-11703-64
14	112	60.2	452	1	US-07-618-312A-2
15	112	60.2	452	1	US-07-618-312A-4
16	112	60.2	452	1	US-08-110-786A-8
17	112	60.2	452	1	US-08-280-228-2
18	112	60.2	452	1	US-08-280-228-4
19	112	60.2	618	1	US-08-668-381A-5
20	107	57.5	22	1	US-08-446-692-5
21	107	57.5	22	2	US-08-488-351A-5
22	107	57.5	22	3	US-09-100-409A-41
23	107	57.5	22	5	PCT-US95-13841-8
24	100	53.8	19	1	US-07-610-525-2
25	94.5	50.8	20	1	US-08-319-704-11
26	88	47.3	17	1	US-07-610-525-3
27	87.5	47.0	22	2	US-08-817-933A-8

```

28 79 42.5 14 1 US-07-610-525-6 Sequence 6, Appli
29 79 42.5 14 1 US-08-787-547-43 Sequence 43, Appli
30 79 42.5 15 1 US-07-610-525-4 Sequence 4, Appli
31 79 42.5 15 1 US-08-787-547-44 Sequence 44, Appli
32 69 37.1 13 1 US-07-610-525-5 Sequence 5, Appli
33 61 32.8 25 1 US-08-446-692-17 Sequence 17, Appli
34 61 32.8 25 2 US-08-488-351A-17 Sequence 4, Appli
35 59 31.7 49 1 US-08-387-156-4 Sequence 4, Appli
36 59 31.7 49 2 US-08-694-865-4 Sequence 4, Appli
37 59 31.7 49 3 US-08-878-748-4 Sequence 4, Appli
38 59 31.7 49 3 US-09-124-491-4 Sequence 4, Appli
39 59 31.7 438 1 US-08-480-604A-23 Sequence 23, Appli
40 59 31.7 438 2 US-08-405-496A-23 Sequence 23, Appli
41 59 31.7 438 4 US-08-915-136-23 Sequence 23, Appli
42 59 31.7 462 1 US-08-480-604A-26 Sequence 26, Appli
43 59 31.7 462 2 US-08-405-496A-26 Sequence 26, Appli
44 59 31.7 462 4 US-08-915-136-26 Sequence 26, Appli
45 59 31.7 544 1 US-08-387-156-10 Sequence 10, Appli

```

ALIGNMENTS

RESULT 1

```

US-08-446-692-14
; Sequence 14, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulants for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-14

```

Query Match 81.7%; Score 152; DB 1; Length 32;
Best Local Similarity 87.9%; Pred. No. 7.7e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

```

QY 1 FNNFTVSFWLRPKVSASHLEGPSLHWSYGLR 33
Db 3 FNNFTVSFWLRPKVSASHLE----HWSYGLR 31

```


QY 1 QYFKANS---KFTGITELG 16
:||: || ||: || ||||
Db 51 EYIRFNSTVGKFGVYTELG 69

RESULT 13
T39087
Hypothetical protein SPAC7D4.09c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39087
R:Gentiles, S.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21826
A:Accession: T39087
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-274 <GEN>
A:Cross-references: EMBL:Z99532; PIDN:CAB16726.2; GSPDB:GN00066; SPDB:SPAC7D4.09c
A:Experimental source: strain 972h-; cosmid c7D4
C:Genetics:
A:Gene: SPDB:SPAC7D4.09c
A:Map position: 1

```

Query Match      33.8%; Score 50; DB 2; Length 274;
Best Local Similarity 29.4%; Pred. No. 12;
Matches 10; Conservative 7; Mismatches 9; Indels 8; Gaps 1;

QY      2 YIKANSKFIGITGLPSLHW-----SYGLRP 27
      | : : | : | | | : : | | |
Db      18 YFTSTILFVSLKNAPSLSWLMKYGGHDNFGLRP 51

RESULT 14
A83487
probable MFS transporter PA1262 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83487
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
N.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; UID:20437337

```

A:Accession: A83487
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <STO>
A:Cross-references: GB:AE004656; GB:AE004091; PIDN:g9947194; PIDN:AAG04651.1; GSPDB:GN001
A:Experimental source: strain PAOL
C:Genetics:
A:Gene: PAL262
C:Superfamily: lincomycin-resistance protein lmrB

	Query Match	33.8%;	Score 50;	DB 2;	Length 480;
Best Local Similarity	88.9%;	Pred. No. 22;			
Matches	8;	Conservative	1;	Mismatches	0;
				Indels	0;
				Gaps	0;

QY	15	LGPSLHWSY	23
		:	
Db	224	LGPSLHWSW	232

RESULT 15
T38962
hypothetical protein SPAC5D6.04 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38962
R:Skelton, J.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1997

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27

|||||||

Db 2 HWSYGLRP 9

RESULT 8

I51423 gonadoliberin precursor - African clawed frog

N;Alternate names: luteinizing hormone releasing hormone

C;Species: Xenopus laevis (African clawed frog)

C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999

C;Accession: I51423

R;Hayes, W.P.; Wray, S.; Battey, J.F.

Endocrinology 134, 1835-1845, 1994

A;Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain

A;Reference number: I51423; MUID:94185563

A;Accession: I51423

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-89 <HAY>

A;Cross-references: GB:L28040; NID:g496291; PIDN:AAA49728.1; PID:g496292

C;Genetics:

A;Gene: GnRH-I

C;Superfamily: gonadoliberin

Query Match

Best Local Similarity 35.1%; Score 52; DB 2; Length 89;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27

|||||||

Db 25 HWSYGLRP 32

RESULT 9

AD3098

periplasmic nitrate reductase large subunit [Imported] - Agrobacterium tumefaciens (strain

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C;Accession: AD3098

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.

; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193

A;Accession: AD3098

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-824 <KUR>

A;Cross-references: GB:AE008689; PIDN:AAL45202.1; PID:g17742881; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: napA

A;Map position: linear chromosome

Query Match

Best Local Similarity 34.8%; Score 51.5; DB 2; Length 824;

Matches 10; Conservative 9; Mismatches 1; Indels 7; Gaps 2;

QY 2 YIKANSKFI-GITELGSLHWSYGLRP 27

|||||

Db 279 FVRNHTKVRGVTDIG-----YGLRP 299

RESULT 10

E98188

periplasmic nitrate reductase precursor napA (AF040988) [imported] - Agrobacterium tumef

C;Species: Agrobacterium tumefaciens

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C;Accession: E98188

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A;Reference number: A97359; PMID:11743194

A;Accession: E98188

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-834 <KUR>

A;Cross-references: GB:AE007870; PIDN:AAK89031.1; PID:g15158825; GSPDB:GN00170

C;Genetics:

A;Gene: AGR_L_917

A;Map position: linear chromosome

Query Match 34.8%; Score 51.5; DB 2; Length 834;

Best Local Similarity 37.0%; Pred. No. 24;

Matches 10; Conservative 9; Mismatches 1; Indels 7; Gaps 2;

QY 2 YIKANSKFI-GITELGSLHWSYGLRP 27

|||||

Db 289 FVRNHTKVRGVTDIG-----YGLRP 309

RESULT 11

C72710

probable fmu protein APE1098 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C;Accession: C72710

R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A;Reference number: A72450; MUID:99310339

A;Accession: C72710

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-388 <RAW>

A;Cross-references: DDBJ:AF000060; NID:g5104188; PIDN:BAA80083.1; PID:d1043869; PID:g

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE1098

Query Match

Best Local Similarity 34.5%; Score 51; DB 2; Length 388;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 14 ELGPSLHWSYGLRP 27

|||||

Db 345 ELGRLTWSMGLRP 358

RESULT 12

S29982

class II histocompatibility antigen - Atlantic salmon

C;Species: Salmo salar (Atlantic salmon)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999

C;Accession: S29982

R;Hordvik, I.

submitted to the EMBL Data Library, October 1992

A;Reference number: S29980

A;Accession: S29982

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-244 <HOR>

A;Cross-references: EMBL:X70166; NID:g64369; PID:g64370

C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 34.1%; Score 50.5; DB 2; Length 244;

Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

C:Genetics:
A:Gene: GDB:GNRH; LHRH; GRH
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A:Map position: 8p21-8p11.2
A:Introns: 47/3; 79/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
F:1-23/Domain: amidated carboxyl end; hormone: hypothalamus; placenta; pyroglutamic acid
F:24-92/Product: signal sequence #status predicted <SIG>
F:24-33/Product: gonadoliberin #status predicted <PGN>
F:37-92/Product: gonadoliberin-associated protein #status predicted <MAT>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 36.5%; Score 54; DB 1; Length 92;
Best Local Similarity 90.0%; Pred. No. 0.97;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27
Db 23 SQHWSYGLRP 32
| | | | | | | | | |

RESULT 5
RHRG
N:Alternate names: gonadoliberin-precursor - rat
N:Contents: gonadoliberin; prolactin release-inhibiting factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A40147; B26173; A48410
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
A:Reference number: A40147; MUID:89384661
A:Accession: A40147
A:Molecule type: DNA
A:Residues: 1-92 <BON>
A:Cross-references: GB:M31670; NID:G204447; PIDN:AAA41264.1; PID:G204448
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
A:Reference number: A94090; MUID:86094338
A:Accession: B26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12579; NID:G204445; PIDN:AAA41263.1; PID:G204446
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blaiock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone
A:Reference number: A48410; MUID:93105480
A:Accession: A48410
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <MAI>
A:Cross-references: GB:S50870; NID:G262059; PIDN:AAB24572.1; PID:G262060
A:Experimental source: thymus
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:121083)
C:Genetics:
A:Introns: 47/3; 79/3
C:Function:
A:Description: stimulates pituitary secretion of lutropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
F:1-23/Domain: amidated carboxyl end; hormone: hypothalamus; placenta; pyroglutamic acid; I
F:24-92/Product: signal sequence #status predicted <SIG>
F:24-33/Product: gonadoliberin #status predicted <PGN>
F:37-92/Product: gonadoliberin #status predicted <GLN>
F:24/Modified site: prolactin release-inhibiting factor #status predicted <PIF>
F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted

F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 36.5%; Score 54; DB 1; Length 92;
Best Local Similarity 90.0%; Pred. No. 0.97;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27
Db 23 SQHWSYGLRP 32
| | | | | | | | | |

RESULT 6
RHPGG
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A:Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.1%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
Db 2 HWSYGLRP 9
| | | | | | | |

RESULT 7
RHSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780
A:Molecule type: protein
A:Residues: 1-10 <BUR>
A:Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.1%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.18;

RESULT 3
RHM5G
gonadolibarin precursor - mouse
N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing
N;Contents: gonadolibarin; gonadolibarin-associated protein (GAP)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C;Accession: A47578
R;Hason, A. &; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolic,
Science 234, 1366-1371, 1986
A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible for
A;Reference number: A47578; MUID:87069928
A;Accession: A47578

R;itali, L.; Rousseau, F.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized by the rat hypothalamus
A:Reference number: A90108; PMID:83126573
A:Accession: A90108
A:Molecule type: protein
A:Residues: 24-33 <TAN>
A:Experimental source: placental trophoblasts
R;Leibovitz, D.; Koch, Y.; Pilzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amstutz, P.
FEBS Lett. 346, 203-206, 1994
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by the rat hypothalamus
A:Reference number: S45718; PMID:94283597
A:Contents: annotation; degradation pathway of synthetic hormone

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 16:00:11 ; Search time 6.52137 Seconds
(without alignments)
412.567 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKFIGITELGFLSHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	74	50.0	1315	1 BTCLTN	tentoxylisin (EC 3
2	54	36.5	67	2 I78541	gonadoliberin (EC 3
3	54	36.5	90	1 RHMSG	gonadoliberin prec
4	54	36.5	92	1 RHUG	gonadoliberin prec
5	54	36.5	92	1 RHRTG	gonadoliberin prec
6	52	35.1	10	1 RHPGG	gonadoliberin - pl
7	52	35.1	10	1 RHSHG	gonadoliberin - sh
8	52	35.1	89	2 I51423	gonadoliberin prec
9	51.5	34.8	824	2 AD3098	periplasmic nitrat
10	51.5	34.8	834	2 E98188	periplasmic nitrat
11	51	34.5	388	2 G72710	probable fmu prote
12	50.5	34.1	244	2 S29882	class II histocomp
13	50	33.8	274	2 T39087	hypothetical prote
14	50	33.8	480	2 AD3098	probable MFS trans
15	49	33.1	452	2 T38962	hypothetical prote
16	49	33.1	484	2 E83245	potassium uptake p
17	49	33.1	3133	2 S52093	hemocytin - silkwo
18	48	32.4	10	1 RHA01	gonadoliberin I -
19	48	32.4	92	2 I50844	gonadoliberin I pr
20	48	32.4	318	2 G97588	probable transketol
21	48	32.4	318	2 AD2810	transketolase [imp
22	48	32.4	794	2 AD0294	pyruvate,water dik
23	47.5	32.1	205	2 C90140	conserved hypothet
24	47	31.8	728	2 E83228	hypothetical prote
25	46.5	31.4	193	2 F89967	hypothetical prote
26	46.5	31.4	245	2 S29980	class II histocomp
27	46.5	31.4	388	2 AB2445	hypothetical prote
28	46	31.1	98	2 I50739	gonadotropin-relea
29	46	31.1	171	2 S38237	hypothetical prote

30	46	31.1	186	2 A90167	adenylate cyclase,
31	46	31.1	349	2 E75611	glucosamine--fruct
32	46	31.1	459	2 G82431	C4-dicarboxylate t
33	46	31.1	522	2 T44369	pyruvate,water dik
34	46	31.1	780	2 D75361	phosphoenolpyruvat
35	46	31.1	4056	2 H96599	protein F14J16.10
36	45.5	30.7	322	2 T38399	probable amidohydr
37	45	30.4	256	2 T50616	hypothetical prote
38	45	30.4	258	2 F72052	peptidyl-prolyl ci
39	45	30.4	258	2 B86573	RKBP-type peptidyl
40	45	30.4	368	2 T27432	hypothetical prote
41	45	30.4	397	2 B70815	probable transamin
42	45	30.4	435	2 C89857	conserved hypothet
43	45	30.4	456	2 S55661	hypothetical prote
44	45	30.4	644	2 S46746	hypothetical prote
45	45	30.4	674	2 H72423	alpha-glucuronidas

ALIGNMENTS

RESULT 1

BTCLTN
tentoxylisin (EC 3.4.24.68) precursor - Clostridium tetani
N:Alternate names: tetanus neurotoxin
C:Species: Clostridium tetani
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A25689; A25757; A25194; A60759; S69348; S09364
R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, EMBO J. 5, 2495-2502, 1986
A>Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b
A:Reference number: A25689; MUID:87053814
A:Accession: A25689
A:Molecule type: DNA
A:Residues: 1-1315 <EIS>
A:Cross-references: GB:X04436; NID:940769; PIDN:CRA28033.1; PID:940770
R:Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A>Title: The complete nucleotide sequence of tetanus toxin.
A:Reference number: A25757; MUID:87040747
A:Accession: A25757
A:Molecule type: DNA
A:Residues: 1-1315 <FA1>
A:Cross-references: GB:X06214; NID:940773; PIDN:CRA29564.1; PID:940774
R:Experimental source: strain CN3911
R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A>Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C i
A:Reference number: A25194; MUID:86085672
A:Accession: A25194
A:Molecule type: DNA
A:Residues: 743-1315 <FA2>
A:Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
A:Accession: B25194
A:Molecule type: protein
A:Residues: 865-894 <FA3>
R:Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A>Title: Isolation, purification, and characterization of fragment B, the NH-2-termin
A:Reference number: A60759; MUID:90035436
A:Accession: A60759
A:Molecule type: protein
A:Residues: 461-475 <MAT>
R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A>Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A:Reference number: JS0098; MUID:89033918
A:Contents: annotation; epitope region
R:Schiavo, G.; Benfenati, F.; Poulin, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, Nature 359, 832-835, 1992
A>Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A:Reference number: S27125; MUID:93063293
A:Contents: annotation

CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF033346; AAB87688.1; -
CC InterPro; IPR002012; GnRH.
CC PRINTS; PR01541; GONADOLIBERN1.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 92 GONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT MOD_RES 24 24 ACTIVITY (BY SIMILARITY).
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 33 33 SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT MOD_RES 33 33 SIMILARITY).
SQ SEQUENCE 92 AA; 10279 MW; ACF74613F456D663 CRC64;

Query Match 32.4%; Score 48; DB 1; Length 92;
Best Local Similarity 57.1%; Pred. No. 2.5;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 14 ELGPSLHWSYGLRP 27
Db 19 ENGSGQWSYGVPR 32
| | :||||:|
| | :||||:|

RESULT 15
GONI_CHICK
ID GONI_CHICK STANDARD; PRT; 92 AA.
AC P37042; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN;
RX MEDLINE=94059355; PubMed=7902095;
RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
RT "Characterization of the chicken pregonadotrophin-releasing
RT hormone-I gene.";
RL J. Mol. Endocrinol. 11:19-29(1993).
RN [2]
RP SEQUENCE OF 24-33.
RC TISSUE=Hypothalamus;
RX MEDLINE=82265778; PubMed=7050119;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RT hormone. II. Isolation and characterization.";
RL J. Biol. Chem. 257:10729-10732(1982).
RN [3]
RP SEQUENCE OF 24-33.
RC TISSUE=Hypothalamus;
RX MEDLINE=82265777; PubMed=7050118;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RT hormone. I. Structural determination on partially purified
RT material.";
RL J. Biol. Chem. 257:10722-10728(1982).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69491; CAA49246.1; -
CC PIR; S33507; S33507.
CC InterPro; IPR002012; GnRH.
CC PRINTS; PR01541; GONADOLIBERN1.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23 GONADOLIBERIN I.
FT CHAIN 24 92 GONADOLIBERIN I.
FT PEPTIDE 24 33 GnRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10206 MW; 61AEB7EBAF508B6A CRC64;

Query Match 32.4%; Score 48; DB 1; Length 92;
Best Local Similarity 87.5%; Pred. No. 2.5;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
Db 25 HWSYGLQP 32
| | | | | | | |
| | | | | | | |

Search completed: October 10, 2002, 16:06:30
Job time : 4.64957 secs

RL S. Afr. J. Sci. 78:124-125(1982).
RN [4]
RP SYNTHESIS OF 24-33.
RX MEDLINE=82265777; PubMed=7050118;
RA King J.A., Millar R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RT hormone. I. Structural determination on partially purified
RT material.";
RL J. Biol. Chem. 257:10722-10728(1982).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69491; CAA49246.1; -
CC PIR; S33507; S33507.
CC InterPro; IPR002012; GnRH.
CC PRINTS; PR01541; GONADOLIBERN1.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23 GONADOLIBERIN I.
FT CHAIN 24 92 GONADOLIBERIN I.
FT PEPTIDE 24 33 GnRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10206 MW; 61AEB7EBAF508B6A CRC64;

Query Match 32.4%; Score 48; DB 1; Length 92;
Best Local Similarity 87.5%; Pred. No. 2.5;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
Db 25 HWSYGLQP 32
| | | | | | | |
| | | | | | | |

Search completed: October 10, 2002, 16:06:30
Job time : 4.64957 secs

DR PROSITE; PS01185; CTCK_1; 1.
DR DR PROSITE; PS01208; WFEC; FALSE_NEG.
DR DR PROSITE; PS01225; CTCK_2; 1.
DR DR PROSITE; PS01285; FA58C_1; 2.
DR DR PROSITE; PS01286; FA58C_2; 2.
KW Lectin; Glycoprotein; Signal; Repeat; cell adhesion.
FT SIGNAL ? 3133
FT CHAIN ? 3133
FT DOMAIN 29 131
FT DOMAIN 153 240
FT DOMAIN 248 613
FT DOMAIN 940 1095
FT DOMAIN 1116 1254
FT DOMAIN 1283 1356
FT DOMAIN 1620 1951
FT DOMAIN 1952 2315
FT DOMAIN 2230 2321
FT DOMAIN 2335 2361
FT DOMAIN 2435 2469
FT DOMAIN 2553 2622
FT DOMAIN 2842 2907
FT DOMAIN 2971 3076
FT DOMAIN 895 914
FT DOMAIN 1267 1270
FT DOMAIN 1425 1428
FT DOMAIN 1447 1450
FT DOMAIN 1474 1479
FT DOMAIN 2148 2153
FT DOMAIN 2156 2159
FT DOMAIN 2341 2344
FT DISULFID 940 1095
FT DISULFID 1116 1254
FT DISULFID 2981 3040
FT DISULFID 2991 3054
FT DISULFID 3004 3070
FT DISULFID 3020 3072
FT DISULFID ? 3075
FT CARBOHYD 151 151
FT CARBOHYD 237 237
FT CARBOHYD 564 564
FT CARBOHYD 1170 1170
FT CARBOHYD 1387 1387
FT CARBOHYD 1622 1622
FT CARBOHYD 1727 1727
FT CARBOHYD 1847 1847
FT CARBOHYD 1975 1975
FT CARBOHYD 1985 1985
FT CARBOHYD 2093 2093
FT CARBOHYD 2113 2113
FT CARBOHYD 2161 2161
FT CARBOHYD 2276 2276
FT CARBOHYD 2451 2451
FT CARBOHYD 2647 2647
FT CARBOHYD 2654 2654
FT CARBOHYD 2663 2663
FT CARBOHYD 2794 2794
FT CARBOHYD 2810 2810
FT CARBOHYD 2865 2865
FT CARBOHYD 2929 2929
FT CARBOHYD 2964 2964
FT CARBOHYD 3028 3028
FT VARIANT 1288 1288
FT VARIANT 1305 1305
FT SEQUENCE 3133 AA; 343350 MW; E5210D5D14A7B2B2 CRC64;
Query Match 33.1%; Score 49; DB 1; Length 3133;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 9 FIGITELGPSLHWSYGLR 26
1: 1111111111
Db 346 FLDVPSLGLMSLQWRGLR 363

RESULT 13
GONI_ALLMI
ID GONI_ALLMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I) (Luliberin I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M., Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT "Primary structure of two forms of gonadotropin-releasing hormone from brains of the American alligator (Alligator mississippiensis)."; Regul. Pept. 33:105-116(1991).
RL CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR; A60066; RHAQ1.
DR InterPro; IPR002012; GNRH.
DR Pfam; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
FT SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;
SQ
Query Match 32.4%; Score 48; DB 1; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.28;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 20 HWSYGLRP 27
11111111
Db 2 HWSYGLQP 9
RESULT 14
GONI_CAVPO
ID GONI_CAVPO STANDARD; PRT; 92 AA.
AC O54713;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I) (Luteinizing hormone releasing hormone I) (Gonadotropin releasing hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].
GN GNRH1 OR GNRH OR LHRH.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY WHITE; TISSUE=Hypothalamus;
RX MEDLINE=97462693; PubMed=9322920;
RA Jimenez-Linan M., Rubin B.S., King J.C.;
RT "Examination of guinea pig luteinizing hormone-releasing hormone gene reveals a unique decapeptide and existence of two transcripts in the brain."; Endocrinology 138:4123-4130(1997).
RL CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.

```

DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 92 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT ACT_SITE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 26 26 ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 33 33 SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT MOD_RES 33 33 SIMILARITY).
SQ SEQUENCE 92 AA; 10197 MW; 4FDBF2C36CF5F63B CRC64;

Query Match 35.1%; Score 52; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
Db 25 HWSYGLRP 32
|||||

RESULT 11
IDQ4_SCHPO STANDARD; PRT; 452 AA.
AC O14197;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 49.6 kDa protein C5D6.04 in chromosome I.
DS SPAC5D6.04.
GS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
ON NCBI_TaxID=4896;
RX STRAIN=972;
RP SEQUENCE FROM N.A.
RA Skellton J., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO YEAST YBR287W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z98056; CAB10852.1; -
DR KW Hypothetical protein; Transmembrane.
FT TRANSMEM 18 38 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 354 374 POTENTIAL.
FT TRANSMEM 390 410 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
SQ SEQUENCE 452 AA; 49575 MW; 71B77EA5725C69A8 CRC64;

Query Match 33.1%; Score 49; DB 1; Length 452;
Best Local Similarity 57.1%; Pred. No. 8.6;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 13 TELGPSLHWSYGLR 26
Db 166 SQLGQALRWSYGYR 179
|||||

```

```

RESULT 12
HMCT_BOMMO STANDARD; PRT; 3133 AA.
ID HMCT_BOMMO
AC P98092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hemocytin precursor (Humoral lectin).
DE Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
ON NCBI_TaxID=7091;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PUYOU X TOKAI; TISSUE=Hemocyte;
RX MEDLINE=95178544; PubMed=7873598;
RA Kotani E., Yamakawa M., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
RA Matsubara F., Tani K., Kadohara K., Kato Y., Mori H.;
RT "Cloning and expression of the gene of hemocytin, an insect humoral
RT lectin which is homologous with the mammalian von Willebrand
RT factor.";
RL Biochim. Biophys. Acta 1260:245-258(1995).
RN [2]
RP SEQUENCE OF 2221-3133 FROM N.A.
RA Kotani E., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
RA Matsubara F., Yamakawa M.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ADHESIVE PROTEIN AND RELATES TO HEMOSTASIS OR
CC ENCAPSULATION OF FOREIGN SUBSTANCES FOR SELF-DEFENSE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN HEMOCYTES DURING LARVAL-PUPAL
CC METAMORPHOSIS.
CC -1- INDUCTION: HEMAGGLUTINATION ACTIVITY IS INCREASED BY BACTERIAL
CC OR VIRAL INFECTION AND INHIBITED BY D-MANNOSE, N-ACETYL-D-
CC GALACTOSAMINE AND D-MALTOSE.
CC -1- PTM: MAY BE CONVERTED INTO THE 260 kDa MATURE HEMOCYTIN BY
CC PROTEOLYSIS.
CC -1- SIMILARITY: TO MAMMALIAN VON WILLEBRAND FACTOR.
CC -1- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT
CC OF HUMAN MUCIN 2.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D29738; BAA06160.1; -
DR EMBL; D14035; BAA03124.1; -
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR000421; F58_C.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR002919; TIL.
DR InterPro; IPR001007; VWFC.
DR InterPro; IPR001846; Vwd.
DR InterPro; IPR001878; Znf_CCHC.
DR InterPro; IPR000083; fibronectin_type_1.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR Pfam; PF01826; TIL; 6.
DR Pfam; PF00094; vwd; 3.
DR Pfam; SM00041; C1; 1.
DR SMART; SM00231; FA58C; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00214; VWC; 3.
DR SMART; SM00216; VWD; 3.
DR SMART; SM00343; Znf_C2HC; 1.

```

```
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U91938; AAB51302.1; -
CC InterPro: IPR002012; GNRH.
CC Pfam: PF00446; GNRH; 1.
CC PROSITE: PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Placenta.
CC NON_TER 1 1 PROGNADOLIBERIN I.
CC CHAIN 1 >63 GONADOLIBERIN I.
CC PEPTIDE 1 10 GNRH-ASSOCIATED PEPTIDE I (BY
CC PEPTIDE 14 >63 SIMILARITY).
CC ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC ACTIVITY (BY SIMILARITY).
CC MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
CC MOD_RES 10 10 SIMILARITY).
CC MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
CC NON_TER 63 63 SIMILARITY).
CC SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;
CC -----
CC Query Match 35.1%; Score 52; DB 1; Length 63;
CC Best Local Similarity 100.0%; Pred. No. 0.45;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 20 HWSYGLRP 27
CC Db 2 HWSYGLRP 9
CC -----
CC RESULT 9
CC GONL_XENLA
CC ID GONL_XENLA STANDARD; PRT; 89 AA.
CC AC P45656;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
CC DE (LR-RH) (Luliberin I).
CC OS Xenopus laevis (African clawed frog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
CC Xenopodinae; Xenopus.
CC NCBI_TaxID=8355;
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE-Forebrain;
CC MEDLINE=94185563; PubMed=8137750;
CC RA Hayes W.P., Wray S., Battey J.F.;
CC RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
CC RT mammalian-like expression pattern and conserved domains in
CC RT GNRH-associated peptide, but brain onset is delayed until
CC RT metamorphosis.";
CC RL Endocrinology 134:1835-1844(1994).
CC CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC CC -1- SUBCELLULAR LOCATION: Secreted.
CC CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L28040; AAA49728.1; -
CC InterPro: IPR002012; GNRH.
CC InterPro: IPR004079; GonadoliberinI.
CC Pfam: PF00446; GNRH; 1.
CC PRINTS; PR01541; GONADOLIBRNI.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal.
CC SIGNAL 1 23 PROGNADOLIBERIN I.
CC CHAIN 24 89 GONADOLIBERIN I.
CC PEPTIDE 24 33 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
CC PEPTIDE 37 89 PEPTIDE.
CC FT PEPTIDE 37 85 GNRH-ASSOCIATED PEPTIDE I (GAP).
CC FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
CC SQ SEQUENCE 89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;
CC -----
CC Query Match 35.1%; Score 52; DB 1; Length 89;
CC Best Local Similarity 100.0%; Pred. No. 0.63;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 20 HWSYGLRP 27
CC Db 25 HWSYGLRP 32
CC -----
CC RESULT 10
CC GONL_TUPGB
CC ID GONL_TUPGB STANDARD; PRT; 92 AA.
CC AC Q95335;
CC DT 15-DEC-1998 (Rel. 37, Created)
CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DE Progonadoliberin I precursor (Contains: Gonadoliberin I (LHRH I)
CC DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
CC DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I].
CC GN GNRH1 OR GNRH.
CC OS Tupia glis belangeri (Common tree shrew).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
CC NCBI_TaxID=9396;
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE-Hypothalamus;
CC MEDLINE=97079639; PubMed=8921350;
CC RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
CC Fernald R.D.;
CC RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
CC RT first direct evidence for mesencephalic GNRH gene expression in a
CC RT placental mammal.";
CC RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC CC HORMONES.
CC CC -1- SUBCELLULAR LOCATION: Secreted.
CC CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U63326; AAB16837.1; -
CC InterPro: IPR002012; GNRH.
CC InterPro: IPR004079; GonadoliberinI.
CC Pfam: PF00446; GNRH; 1.
CC PRINTS; PR01541; GONADOLIBRNI.
CC DR
```

CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S50870; AAB24572.1; -.
 CC EMBL; M12579; AAA41263.1; -.
 CC EMBL; M31670; AAA41264.1; -.
 CC EMBL; M15527; AAA42141.1; ALT_SEQ.
 CC EMBL; M15529; AAA42139.1; -.
 CC EMBL; M15528; -; NOT_ANNOTATED_CDS.
 CC PIR; B26173; RHRTG.
 CC PIR; A48410; A48410.
 CC InterPro: IPR002012; Gnrh.
 CC InterPro: IPR004079; GonadoliberinI.
 CC Pfam; PF00446; Gnrh; 1.
 CC PRINTS; PR01541; GONADOLIBRNI.
 CC PROSITE; PS00473; GNRH; 1.
 CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 CC Placenta; Signal.
 CC FT SIGNAL 1 23 PROGNADOLIBERIN I.
 CC FT CHAIN 24 92 GONADOLIBERIN I.
 CC FT PEPTIDE 24 33 PROLACTIN RELEASE-INHIBITING FACTOR I.
 CC FT PEPTIDE 37 92 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 CC FT ACT_SITE 26 26 ACTIVITY.
 CC FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 CC SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
 CC -----
 CC Query Match 36.5%; Score 54; DB 1; Length 92;
 CC Best Local Similarity 90.0%; Pred. No. 0.33;
 CC Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC QY 18 SLHWSYGLRP 27
 CC I | | | | | | | |
 CC Db 23 SQHWSYGLRP 32
 CC -----
 CC RESULT 7
 CC GONI_SHEEP STANDARD; PRT; 61 AA.
 CC ID GONI_SHEEP STANDARD; PRT; 61 AA.
 CC AC Q28588;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 CC (luteinizing hormone releasing hormone I) (Gonadotropin releasing
 CC hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I]
 CC DE (Fragment).
 CC GN GNRHI OR GNRH OR LHRH.
 CC OS Ovis aries (Sheep).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC OC Bovidae; Caprinae; Ovis.
 CC OX NCBI_TaxID=9940;
 CC RN [1]
 CC RP SEQUENCE OF 12-61 FROM N.A.
 CC RC STRAIN=WESTERN RANGE; TISSUE=Hypothalamus;
 CC RA Rodriguez R.E., Wise M.E.;
 CC RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 CC RN [2]
 CC RP SEQUENCE OF 1-10.

RX MEDLINE-72094314; PubMed-4550508;
 RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
 RA Fellows R., Blackwell R., Vale W., Guillemin R.;
 RT "Primary structure of the ovine hypothalamic luteinizing hormone-mass
 RT releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
 RT spectrometry-decapeptide-Edman degradation)";
 RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U02517; AAA03433.1; -.
 CC PIR; A93780; RHSHG.
 CC InterPro: IPR002012; Gnrh.
 CC Pfam; PF00446; Gnrh; 1.
 CC PROSITE; PS00473; GNRH; 1.
 CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 CC Placenta.
 CC FT NON_TER 1 1 PROGNADOLIBERIN I.
 CC FT CHAIN 1 >61 GONADOLIBERIN I.
 CC FT PEPTIDE 1 10 GNRH-ASSOCIATED PEPTIDE I.
 CC FT PEPTIDE 14 >61 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 CC FT ACT_SITE 3 3 ACTIVITY.
 CC FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
 CC FT NON_TER 61 61
 CC FT NON_TER 61 61
 CC SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;
 CC -----
 CC Query Match 35.1%; Score 52; DB 1; Length 61;
 CC Best Local Similarity 100.0%; Pred. No. 0.43;
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 20 HWSYGLRP 27
 CC I | | | | | | | |
 CC Db 2 HWSYGLRP 9
 CC -----
 CC RESULT 8
 CC GONI_MESAU STANDARD; PRT; 63 AA.
 CC ID GONI_MESAU STANDARD; PRT; 63 AA.
 CC AC O09163;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 CC (luteinizing hormone releasing hormone I) (Gonadotropin releasing
 CC hormone I) (Gnrh I) (Luliberin I); Gnrh-associated peptide I]
 CC DE (Fragment).
 CC GN GNRHI OR GNRH OR LHRH.
 CC OS Mesocricetus auratus (Golden hamster).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC OC Mesocricetus.
 CC OX NCBI_TaxID=10036;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
 CC RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -!- SUBCELLULAR LOCATION: Secreted.


```

Db      23 SQHWSYGLRP 32

RESULT 5
GONL_HUMAN
ID      GONL_HUMAN  STANDARD;      PRT;      92 AA.
AC      P01148;
DT      21-JUL-1986 (Rel. 01, Created)
DT      01-APR-1988 (Rel. 07, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Progonadolibirin I precursor [Contains: Gonadolibirin I (LHRH I)
DE      (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE      hormone I) (Gnrh I) (Luliberin I) (Gonadorelin); Gnrh-associated
DE      peptide I].
DE      GNRH1 OR GNRH OR LHRH.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89366682; PubMed=2671939;
RA      Haylick J.S., Adelman J.P., Seeburg P.H.;
RT      "The complete nucleotide sequence of the human gonadotropin-releasing
RT      hormone gene.";
RL      Nucleic Acids Res. 17:6403-6403(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86094338; PubMed=2867548;
RA      Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT      "Isolation of the gene and hypothalamic cDNA for the common precursor
RT      of gonadotropin-releasing hormone and prolactin release-inhibiting
RT      factor in human and rat.";
RL      Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=85012739; PubMed=6090951;
RA      Seeburg P.H., Adelman J.P.;
RT      "Characterization of cDNA for precursor of human luteinizing hormone
RT      releasing hormone.";
RL      Nature 311:666-668(1984).
RN      [4]
RP      SEQUENCE OF 24-33.
RX      MEDLINE=83126573; PubMed=6760865;
RA      Tan L., Rousseau P.;
RT      "The chemical identity of the immunoreactive LHRH-like peptide
RT      biosynthesized in the human placenta.";
RL      Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC      -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC      THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC      HORMONES.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- PHARMACEUTICAL: Available under the names Factrel (Averst Labs),
CC      Lutrepulse or Lutrelief (ferring Pharmaceuticals) and Relisorm
CC      (Serono).
CC      -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X01059; CAA25526.1; -
DR      EMBL; M12578; AAA35916.1; -
DR      EMBL; X15215; CAA33285.1; -
DR      PIR; A01410; RHUG.
DR      PIR; A26173; A26173.
DR      PIR; S05308; S05308.
DR      MIM; 152760; -
DR      InterPro; IPR002012; Gnrh.

DR      InterPro; IPR004079; GonadolibirinI.
DR      Pfam; PF00446; Gnrh; 1.
DR      PRINTS; PR01541; GONADOLIBRNI.
DR      PROSITE; PS00473; GNRH; 1.
KW      Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW      Placenta; Pharmaceutical; Signal.
FT      SIGNAL          1      23
FT      CHAIN           24      92
FT      PEPTIDE         24      33
FT      PEPTIDE         37      92
FT      ACT_SITE        26      26
FT      MOD_RES         24      24
FT      MOD_RES         33      33
FT      CONFLICT        16      16
FT      CONFLICT        W -> S (IN REF. 3).
SQ      SEQUENCE      92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match          36.5%; Score 54; DB 1; Length 92;
Best Local Similarity 90.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      18 SLHWSYGLRP 27
Db      23 SQHWSYGLRP 32
      | | | | | | | |
      | | | | | | | |

RESULT 6
GONL_RAT
ID      GONL_RAT  STANDARD;      PRT;      92 AA.
AC      P07490;
DT      01-APR-1988 (Rel. 07, Created)
DT      01-APR-1988 (Rel. 07, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Progonadolibirin I precursor [Contains: Gonadolibirin I (LHRH I)
DE      (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE      hormone I) (Gnrh I) (Luliberin I); Prolactin release-inhibiting factor
DE      I].
DE      GNRH1 OR GNRH.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86094338; PubMed=2867548;
RA      Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT      "Isolation of the gene and hypothalamic cDNA for the common precursor
RT      of gonadotropin-releasing hormone and prolactin release-inhibiting
RT      factor in human and rat.";
RL      Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89384661; PubMed=2476669;
RA      Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT      "The rat gonadotropin-releasing hormone: SH locus: structure and
RT      hypothalamic expression.";
RL      Mol. Endocrinol. 3:1257-1262(1989).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93103480; PubMed=1468115;
RA      Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT      "Thymocytes express a mRNA that is identical to hypothalamic
RT      luteinizing hormone-releasing hormone mRNA.";
RL      Cell. Mol. Neurobiol. 12:447-454(1992).
RN      [4]
RP      SEQUENCE OF 1-47 FROM N.A.
RX      TISSUE=Heart;
RX      MEDLINE=87149087; PubMed=3547652;
RA      Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT      "Two mammalian genes transcribed from opposite strands of the same
RT      DNA locus.";
RL      Science 235:1514-1517(1987).

```

RESULT 3

GONI_MOUSE STANDARD; PRT; 90 AA.
 AC P13562;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
 DE I].
 DE GNRH1 OR GNRH.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=67069928; PubMed=3024317;
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
 RA Phillips H.S., Nikolic K., Seeburg P.H.;
 RT "A deletion truncating the gonadotropin-releasing hormone gene is
 RT responsible for hypogonadism in the hpg mouse.";
 RL Science 234:1366-1371(1986).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M14872; AAA37717.1; -.
 DR MGD; MGI:95789; GnRh.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBRNI.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 90 PROGONADOLIBERIN I.
 FT PEPTIDE 22 31 GONADOLIBERIN I.
 FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
 FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 36.5%; Score 54; DB 1; Length 90;
 Best Local Similarity 90.0%; Pred. No. 0.32;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27
 | | | | | | | | | |
 Db 21 SQHWSYGLRP 30

RESULT 4

GONI_PIG STANDARD; PRT; 91 AA.
 AC P49921;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
 DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
 DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
 GN GNRH1 OR GNRH.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=72114303; PubMed=4946067;
 RA Baba Y., Matsuo H., Schally A.V.;
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.
 RT Confirmation of the proposed structure by conventional sequential
 RT analyses.";
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
 CC [3]
 CC SYNTHESIS OF GONADOLIBERIN.
 CC MEDLINE=72065376; PubMed=4942726;
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 RT phase method.";
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
 CC [4]
 CC SYNTHESIS OF GONADOLIBERIN.
 CC MEDLINE=72117544; PubMed=4946275;
 RA Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L32864; AAA31066.1; -.
 DR PIR; A01411; RHPGG.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBRNI.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 91 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

Query Match 36.5%; Score 54; DB 1; Length 91;
 Best Local Similarity 90.0%; Pred. No. 0.32;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27
 | | | | | | | | | |

EMBO J. 11:3577-3583(1992).
[7]
IDENTIFICATION OF SUBSTRATE.
MEDLINE=93063293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.,
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin.";
RL Nature 359:832-835(1992).
[8]
X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX MEDLINE=97475217; PubMed=9334741;
RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA Sax M.;
RT "Structure of the receptor binding fragment HC of tetanus
RT neurotoxin.";
RL Nat. Struct. Biol. 4:788-792(1997).
CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
CC BOND OF SYNAPTOSOMAL-2.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-|-PHE-77 BOND IN
CC SYNAPTOSOMAL-2.
CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC AND ARE NON-TOXIC AFTER SEPARATION.
CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC GANGLIOSIDE RECEPTORS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X04436; CAA28033.1; -;
DR EMBL; M12739; AAA23282.1; -;
DR EMBL; X06214; CAA29564.1; -;
DR PIR; A25689; BTCLTN.
DR PDB; 1AF9; 29-APR-98.
DR PDB; 1A8D; 14-OCT-98.
DR MEROPS; M27.001; -;
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_Mtpeptidse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
KW 3D-structure.
FT INIT_MET 0
FT CHAIN 0
FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 233 233 BY SIMILARITY.
FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 669 689 POTENTIAL.
FT DISULFID 438 466 INTERCHAIN.
FT DISULFID 1076 1092
SQ SEQUENCE 1314 AA; 150550 MW; 134C3657133BF81D CRC64;

Query Match 50.0%; Score 74; DB 1; Length 1314;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITEL 15

|||||
Db 829 QYKANSKFIGITEL 843

RESULT 2
GONI_MACMU STANDARD; PRT; 67 AA.
ID GONI_MACMU
AC P52247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE Luteinizing hormone releasing hormone I] (Gonadotropin releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
GN GNRL1 OR GnRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RL Neuroendocrinology 60:346-359(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; S75918; AAB33096.1; -;
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 5 BY SIMILARITY.
FT CHAIN 6 >67 PROGNADOLIBERIN I.
FT PEPTIDE 6 15 GONADOLIBERIN I.
FT PEPTIDE 19 >67 GnRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
FT NON_TER 67 67
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 36.5%; Score 54; DB 1; Length 67;
Best Local Similarity 90.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 18 SLHWSYGLRP 27
I |||||
Db 5 SQHWSYGLRP 14

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 15:58:41 ; Search time 3.64957 Seconds
(without alignments)
297.061 Million cell updates/sec

Title: US-09-848-834A-11
Perfect score: 148
Sequence: 1 QYIKANSKFIGITELGSPSLHWSYGLRFX 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	74	50.0	1314	1	TETX_CLOTE
2	54	36.5	67	1	GON1_MACMU
3	54	36.5	90	1	GON1_MOUSE
4	54	36.5	91	1	GON1_PIG
5	54	36.5	92	1	GON1_HUMAN
6	54	36.5	92	1	GON1_RAT
7	52	35.1	61	1	GON1_SHEEP
8	52	35.1	63	1	GON1_MESAU
9	52	35.1	89	1	GON1_XENLA
10	52	35.1	92	1	GON1_TURPB
11	49	33.1	452	1	YDQ4_SCHPO
12	49	33.1	3133	1	HMCT_BOOMO
13	48	32.4	10	1	GON1_ALLMI
14	48	32.4	92	1	GON1_CAVPO
15	48	32.4	92	1	GON1_CHICK
16	46	31.1	94	1	GON1_HAPBU
17	46	31.1	95	1	GON1_MORSA
18	46	31.1	95	1	GON1_PAGMA
19	46	31.1	95	1	GON1_SPARU
20	46	31.1	99	1	GON1_DICLA
21	46	31.1	110	1	YHBJ_ACTAC
22	46	31.1	459	1	DCUC_VIBCH
23	46	31.1	780	1	PPSA_DEIRA
24	45.5	30.7	322	1	YAU8_SCHPO
25	45	30.4	258	1	MTP_CHLPN
26	45	30.4	674	1	YHJ9_YEAST
27	45	30.4	644	1	AGUA_THEMEA
28	45	30.4	831	1	NAPA_ALCEU
29	44	29.7	66	1	VG84_BPMLS
30	44	29.7	90	1	GON8_RANDY
31	44	29.7	255	1	YABD_BACSU
32	44	29.7	256	1	YD83_METJA
33	44	29.7	293	1	CFXQ_GUITH

34	44	29.7	357	1	YQO2_CAEEL	Q09305 caenorhabdi
35	44	29.7	459	1	YGN9_YEAST	P53083 saccharomyc
36	44	29.7	613	1	FIBP_ADEMI	P19721 mouse adeho
37	44	29.7	664	1	ACES_ANOST	P56161 anopheles s
38	44	29.7	1016	1	FDOG_ECOLI	P32176 escherichia
39	43.5	29.4	90	1	GON3_DICLA	Q91a09 dicentrarch
40	43.5	29.4	231	1	CYSH_BACHD	O9kct3 bacillus ha
41	43.5	29.4	236	1	CH22_BACSU	O06737 bacillus su
42	43	29.1	185	1	PTH_RICPR	O9zcv4 rickettsia
43	43	29.1	193	1	MSG1_HUMAN	Q99966 homo sapien
44	43	29.1	229	1	COX2_PARLI	P12701 paracentrot
45	43	29.1	229	1	COX2_PISOC	P25002 pisaster oc

ALIGNMENTS

RESULT 1
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).
OS Clostridium tetani.
OG Plasmid.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_taxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
homology with botulinum toxins.";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE OF 742-1314 FROM N.A.
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
fragment C in Escherichia coli.";
RL J. Bacteriol. 165:21-27(1986).
RN [4]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Kriegstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
in tetanus toxin.";
RL Eur. J. Biochem. 188:39-45(1990).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Kriegstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
identification of cleavage sites.";
RL Eur. J. Biochem. 202:41-51(1991).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010946; PubMed=1396558;
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
neurotransmitter release and protease activity depend on zinc.";